La mayoría de libros de Virus editorial se encuentran bajo licencias libres y para su libre descarga; una apuesta por el acceso libre al conocimiento y la cultura, que consideramos imprescindible en una sociedad en la que las desigualdades sociales también se traducen en desigualdad a la hora de acceder a los contenidos culturales. Pero los proyectos autogestionarios y alternativos, como Virus editorial, suelen tener importantes límites económicos, que en ocasiones afectan a su sostenibilidad o impiden asumir proyectos más costosos o arriesgados. En la medida en que ofrecemos buena parte de nuestro trabajo para lo común, creemos importante crear también formas de colaboración en la sostenibilidad del proyecto:

a) Puedes hacerte soci@ de Virus ingresando un mínimo de 50 € a modo de cuota anual, recibiendo una novedad de tu elección y obteneniendo descuentos en tus compras en nuestra web.

b) Puedes suscribirte a Virus durante un año, aportando 200 €, recibiendo todos los libros de Virus durante 12 meses, dos libros de fondo y descuentos en tus compras en nuestra web.

c) También puedes hacer una donación de cualquier cantidad a través de Paypal.

Google
NOVUS ORDO SECLORUM

Ippolita

El lado oscuro de Google

Historia y futuro de la industria de los metadatos

Virus editorial

Creative Commons
LICENCIA CREATIVE COMMONS
autoría - no derivados - no comercial 1.0

- Esta licencia permite copiar, distribuir, exhibir e interpretar este texto, siempre y cuando se cumplan las siguientes condiciones:
Autoría-atribución: se deberá respetar la autoría del texto y de su traducción. Siempre habrá de constar el nombre del autor/a y del traductor/a.
No comercial: no se puede utilizar este trabajo con fines comerciales.
No derivados: no se puede alterar, transformar, modificar o reconstruir este texto. Los términos de esta licencia deberán constar de una manera clara para cualquier uso o distribución del texto. Estas conciciones sólo se podrán alterar con el permiso expreso del autor/a.

Este libro tiene una licencia Creative Commons Attribution-NoDerivs-NonCommercial. Para consultar las condiciones de esta licencia se puede visitar: http://creativecommons.org/licenses/by-nd-nc/1.0/ o enviar una carta a Creative Commons, 559 Nathan Abbot Way, Stanford, California 94305, EE. UU.

© 2010 de la presente edición, Virus editorial
© 2007 Giangacomo Feltrinelli editore, Milán

Colectivo Ippolita
El lado oscuro de Google
Historia y futuro de la industria de los metadatos

Maquetación: Virus editorial

Cubierta e ilustración de guardas: Pilar Sánchez Molina

Traducción: María Grazia Macchia y Giuseppe Maio

Primera edición: octubre de 2010

Lallevir SL / VIRUS editorial
C/ Aurora, 23 baixos, 08001 Barcelona
T. / Fax: 93 441 38 14
C/e.: virus@pangea.org
www.viruseditorial.net

Impreso en:
Imprenta LUNA
Muelle de la Merced, 3, 2.º izq.
48003 Bilbao
Tel.: 94 416 75 18
Fax.: 94 415 32 98
C/e.: luna@imprentaluna.es

ISBN-13: 978-84-92559-23-7
Depósito legal:

Índice

Introducción 9

I. Historias de motores 17
Motores de búsqueda 19
El nacimiento de Google: en el principio fue el garaje o, más bien, la universidad 23
Google.com: publicidad directa dentro de las páginas 29
El autoservicio publicitario: más allá de la burbuja de la *new economy* 33
Estilo, morfología y multiplicación de los servicios 34
Google, el gigante bueno, cotiza en Bolsa 40
Google, Inc.: el monopolio de la búsqueda 42

II. BeGoogle 49
La huida de cerebros a Google: la guerra por el control de la Web 51
Largas colas en las redes: las entrañas de la economía de la búsqueda, Google versus Microsoft 54
La guerra de los estándares 59
Arma n.º 1: Googleplex, el capitalismo blando 62
Arma n.º 2: perfeccionar la estrategia de la acumulación 67
Arma n.º 3: la imagen lo es todo, pero algo de filosofía no viene mal 69
Arma n.º 4: Google y el open source 74

III. Google open source 77
Teoría y prácticas: *open* no es *free* 79
Los hackers de Stanford 81
La era de la *open source economy* 84
Seducir a los hackers: autonomía, dinero fácil e instrumentos gratuitos 88
Ambientes híbridos entre universidad y empresa 96

IV. Algoritmos, ¡qué pasión! 99
Algoritmos y vida real 101
La estrategia de la objetividad 104
Spiders, base de datos y búsqueda 107
De la *brand-identity* a la interfaz participativa 111

PageRank(™) o la autoridad absoluta de un mundo cerrado 117
PageRank(™) o la moneda de la ciencia 121

V. Además, otras funcionalidades maliciosas 127
Los filtros en los algoritmos: bancos de datos preconfeccionados y control de los usuarios 129
Los *coockie*s de Google, galletas que dejan huella 134
Onanismo tecnológico: crea, recrea y consume tus contenidos 137
Los navegadores como ambientes de desarrollo 141
Privacidad, paranoias y poderes 144

VI. Calidad, cantidad y relación 147
La emergencia de la información 149
Cantidad y calidad 153
El mito de la búsqueda instantánea 154
Detrás del velo del mito 157
Modelos de búsqueda 159

VII. Tecnocracia 163
Tecnocracia: los expertos de la ciencia 166
Milagros de la tecnología: de las opiniones subjetivas a la verdad objetiva 169
Público y privado 171
Líneas de fuga: P2P, medias independientes, criptografía, blogs, FOAF 175

Conclusiones 181

Postfacio. Una historia simple 187

Apéndices 199
I. El fin del mundo en una galleta 199
II. Interzonas: influencias y dominios en los mundos digitales 201

Agradecimientos

Gracias a Schiaffino y Asmagita, la casa de Suna y Villa Andreoli, las chicas y toda la agencia de comunicación xX(y), Rossella Di Leo, Sentia s.r.l., Laboratorio Cotech, Federico Cabitza, Matteo Valsana, Wadada y la casa de Nápoles, la Re, el Pub24, la Scighera, Maria la panadera, enomix y blicero.

Nota de presentación

Ippolita es un grupo de investigación compuesto por hackers y activistas sociales, un servidor y una comunidad de «escribientes», una encrucijada para compartir instrumentos y conocimientos entre el lenguaje del mundo digital y el lenguaje de la escritura. Su trabajo es un modo de acción directa, un llamamiento a no quedarse mirando, sino a intervenir en el torbellino de la tecnocultura, desarrollando formas de autoformación y de puesta en común de competencias a través de prácticas comunitarias de escritura como modelo de autodefensa digital.

Ippolita como colectivo tiene en su haber otro libro, *Open non è Free* (Eleuthera, 2005), y además gestiona un servidor independiente con el fin de dotarse de instrumentos independientes para elaborar proyectos y compartir información.

Ippolita.net quiere ser un mapa en continua evolución y transformación capaz de enlazar universos diferentes, en la búsqueda permanente de todos los enlaces posibles, no solo en la Red. Tanto para elaborar *El lado oscuro de Google* como *Open non è Free* se han valido de Riso, un wiki y CMS (sistema de gestión de contenidos) de creación propia que sirve para proyectos web de escritura colectiva.

Introducción

Google es el motor de búsqueda más conocido y utilizado en todo Internet, hasta tal punto que en los últimos años se ha convertido en el principal medio de acceso a la Red. De forma progresiva, los internautas se han ido adaptando a su interfaz sobria y fiable, a las inserciones publicitarias omnipresentes, pero arrinconadas; han comenzado a disfrutar de sus cómodos servicios y su uso rutinario se ha transformado en una forma de comportamiento: «Si no lo sabes, pregúntale a Google». Se recurre a Google incluso cuando se podría utilizar el *post-it* pegado a la nevera, buscar en la agenda, consultar las Páginas Amarillas u hojear la enciclopedia que está acumulando polvo en las estanterías, junto con las demás obras de consulta en papel, demasiado duras y difíciles de examinar.

Google ha sido capaz de explotar magistralmente nuestra necesidad de sencillez. Google anhela ser el motor de búsqueda perfecto, capaz de entender exactamente las preguntas de los usuarios y devolverles, en un santiamén, exactamente lo que desean. Las inocentes interfaces, que ya pueden ser altamente personalizadas directamente en la Web, y aun así inexorablemente identificables por su estilo minimalista, constituyen una forma de huida cotidiana desde la claustrofobia de las mesas digitales para un número impresionante de usuarios, en constante alza. Son una bocanada de aire fresco,

ventanas privilegiadas abiertas de par en par al fascinante mundo de la Red. ¿Cuántas personas utilizan Google como página inicial de su navegador? Sin embargo, detrás de esa sencillez y facilidad de uso se oculta un coloso, un sistema increíblemente complejo y sutil para la gestión de los conocimientos en el maremágnum de la Red. Google ofrece decenas de servicios gratuitos, para satisfacer todo deseo de búsqueda y comunicación: e-mails, chats, newsgroups, sistemas de indexación de los files en el ordenador, archivos de imágenes, vídeos, libros y mucho más. ¿Por qué? ¿Qué gana? Criticar a Google mediante una ponderación meticulosa de su historia y la deconstrucción de los objetos matemáticos que lo componen representa una ocasión para desvelar una precisa estrategia de dominio cultural. La presente investigación pretende proporcionar un método de análisis más general, útil para el descubrimiento de los tejemanejes de muchas de las aplicaciones que nos hemos acostumbrado a utilizar.

El libro arranca con una breve panorámica sobre la historia de los motores de búsqueda, examina por tanto los momentos más significativos del avance de Google. Tras sobrevivir indemne al estallido de la burbuja de la *new economy*, Google ha entretejido relaciones sólidas con distintas multinacionales de tecnologías de la información (*information tecnology*, IT). La continua expansión de sus actividades en cada sector de las comunicaciones digitales está difundiendo un inconfundible estilo y está modelando un completo universo cultural, el universo de la Web.

Don't be evil (no seas malo) es el lema de Sergey Brin y Larry Page, los dos fundadores de Google. Los ex estudiantes de Stanford, gracias a una meticulosa gestión de su propia imagen, han creado un «gigante bueno», impaciente por archivar nuestras «intenciones de búsqueda» en sus ilimitadas bases de datos. El álter ego digital de millones de usuarios parece estar en buenas manos, confiado al principal *datacenter* de Mountain View, California, conocido como Googleplex. Aquí, como en los demás centros de archivo de datos de

Google —que están surgiendo como setas en todo el mundo— se perfeccionan verdaderas armas para combatir en la guerra por el control de las redes. En primer lugar, se difunde la práctica del capitalismo blando de la abundancia, una estrategia de control biopolítico en sentido estricto, que ofrece ambientes de trabajo confortables, palmaditas en la espalda y gratificaciones a los empleados. Los trabajadores, satisfechos y halagados, se dejan explotar alegremente y se convierten en los mayores sostenedores de la empresa, orgullosos de promocionar una imagen ganadora y «buena».

Los objetivos y los métodos de Google son buenos para todo el mundo; de hecho, la filosofía de la empresa, basada sobre la excelencia de matriz académica, en la apuesta por la innovación y la investigación, se encuentra expresada en diez verdades breves en el propio sitio del motor de búsqueda. Estos diez mandamientos constituyen una especie de buena nueva de la era informática, el Google-pensamiento, divulgado con la ayuda de verdaderos y auténticos «evangelizadores» (*evangelits*), es decir, personalidades punteras del mundo informático. Una última arma, aunque no la menos importante, es la cooptación de las metodologías de desarrollo cooperativo típicas del open source y el uso de softwares libres, no protegidos por copyright o patentes, como base para sus propios productos. De esta forma, Google reduce los costes para la implementación de sus servicios, se asegura el apoyo de técnicos, lumbreras y hackers de todo tipo y se hace pasar por el defensor de la causa de la libre circulación de los saberes, ya que el uso del motor de búsqueda parece ofrecer la mejor forma de acceso gratuito a la Red.

Pero el sueño de Brin y Page de «Google contenedor de todo Internet», cultivado desde los tiempos de la universidad, no es nada más que una idea demagógica, útil para afirmar un culto casi positivista de la objetividad científica: en el caos de la Red solo una técnica superior puede hacerse garante de la transparencia de los procesos, de la exactitud de las respuestas, incluso de la democracia.

De hecho, Google se autodefine como un instrumento «democrático», basado en el presunto carácter «democrático» de la Web. Su algoritmo de indexación de la Red, PageRank(™), se encarga de copiar los datos digitales en los *datacenter*, explotando las conexiones asociadas a cada página singular para determinar su valor. En la práctica, Google interpreta una conexión de la página A a la página B como un voto expresado por la primera sobre la segunda. Pero no se limita a calcular el número de votos, o conexiones, asignadas a una página. Google analiza la página que ha asignado el voto: los votos expresados por páginas «importantes» tienen más relevancia y, por lo tanto, contribuyen a dar «mayor valor» también a las páginas linkadas. El PageRank(™) asigna a los sitios web importantes y de alta calidad un voto más elevado, utilizando filtros y criterios no públicos, que Google tiene en cuenta cada vez que realiza una búsqueda. La «democracia» de Google ordena, pues, la Red en base al número de votos recibidos por cada página y a la importancia de esos votos: una democracia filtrada por la tecnología.

Existen algunos secretos alrededor del coloso de Mountain View, muchos de los que, como veréis, son secretos a voces. El halo de leyenda que envuelve la tecnología googleana se basa en gran parte en la ausencia de una educación de base, de rudimentos prácticos para enfrentarse culturalmente a la onda larga de la revolución tecnológica. Por ejemplo, la extraordinaria rapidez de los frutos de la búsqueda es producto de una cuidadosa selección nada transparente. De hecho, ¿cómo podrían millones de usuarios hojear contemporáneamente en todo instante toda la base de datos de Google si no existieran filtros oportunos para reducir el ámbito de la búsqueda, por ejemplo limitándolo a los datos en su idioma de origen? Y si existen filtros creados para garantizar una mejor navegación lingüística, ¿no es lícito suponer que existan muchos más filtros, estudiados para canalizar incluso las preferencias de los internautas? El prodigio de Google es en realidad una tecnología opaca y encubierta por el copyright y

por acuerdos de no divulgación de sus recursos. La búsqueda no es ni transparente ni democrática, como se nos quiere hacer creer: no podría serlo tanto por razones técnicas como por razones económicas.

El campo blanco de Google en el que se insertan las palabras claves para las búsquedas es una puerta estrecha, un filtro nada transparente, que controla y dirige el acceso a las informaciones. En cuanto mediador de información, un simple motor de búsqueda se hace instrumento de gestión del saber y se halla capacitado para ejercer un poder enorme, convirtiéndose en una autoridad absoluta en un mundo cerrado. El modelo cultural de Google es por lo tanto expresión directa de un dominio tecnocrático.

Con el presente trabajo Ippolita quiere subrayar el problema o, más bien, la urgencia social de alfabetización y orientación crítica del gran público acerca del tema de la gestión de los conocimientos (*knowledges management*). Internet ofrece extraordinarias oportunidades de autoinformación a los usuarios, tantas como para superar incluso la formación universitaria, especialmente en ámbitos como la comunicación y la ingeniería informática. El movimiento del software libre, como ha mostrado Ippolita en otros trabajos, es el ejemplo más evidente de la necesidad de autoformación continua y de la posibilidad de autogestión de los instrumentos digitales.

Sin embargo, hay un reverso de esta moneda doblemente negativo: por un lado, el menosprecio de la formación humanista, que en la Red dispone de pocos y mal organizados ámbitos de referencia; por el otro, el sustancial colapso cognitivo del usuario medio. Desorientados por la redundancia de los datos disponibles en la Red, nos entregamos a aquellos puntos de referencia de mayor visibilidad —de los que Google solo es el ejemplo más patente— sin preguntarnos qué acontece detrás de los bastidores; registramos nuestros datos personales con ligereza, seducidos por el mero uso de servicios decididamente eficaces y, como todavía acostumbra a ser en buena parte de la Red, totalmente gratuitos.

Ippolita intenta subrayar el vacío en la divulgación científica de los fenómenos tecnológicos que afecta al conjunto de la sociedad. Abundan los manuales técnicos, la sociología habla con desenvoltura de sociedades en red, la política se atreve incluso a imaginar una futurible *open society*, en la que las redes serán el sustrato tecnológico de la democracia global. ¿Pero cuántos asiduos internautas saben qué es un algoritmo? Más bien pocos y, sin embargo, la mayoría de ellos confían en la respuesta de PageRank(™), precisamente un algoritmo que selecciona sin pausa los resultados de sus preguntas y dirige su experiencia en la Red. Hace falta el valor de centrarse de nuevo en la divulgación científica, sin encerrarse en la torre de marfil del saber académico. Hace falta hablar de macroeconomía sin ser economistas, de infomediación sin ser expertos en comunicación, de autoinformación sin ser educadores, de autogestión de los instrumentos digitales sin ser políticos. Es preciso provocar debates insistiendo sobre conceptos básicos como «algoritmo», «datos sensibles», «privacidad», «verdad científica», «redes de comunicación», demasiado a menudo debatidos entre la *Authority* y garantes que no pueden garantizar absolutamente nada.

La costumbre de delegar provoca un desinterés generalizado por los grandes cambios en curso en nuestro mundo tecnológico, que acontecen con sordina o encubiertos por el humo mediático, sin haber sido en absoluto asimilados por el gran público.

La postura más común oscila entre el encantamiento maravillado y la frustración frente a los continuos, incomprensibles «milagros de la tecnología»; se llega muy a menudo a la adoración mística, como si lo digital recubriera el mundo de una aura esotérica, comprensible exclusivamente por los pocos iniciados, conjugada con la frustración por la propia incapacidad de oficiar el culto al nuevo progreso de forma adecuada.

Al grupo de investigadores de Ippolita le une precisamente la convicción de que mediante el intercambio y el diálogo

entre habilidades y lenguajes diferentes será posible convertir la llamada Revolución Digital en una materia útil para comprender la contemporaneidad, sus anomalías y probablemente incluso el tiempo por venir. La investigación científica, la tradición humanista, las pasiones políticas, el método feminista son otros lenguajes útiles en esta exploración.

La actividad de Ippolita revela que «poner en común» no es suficiente, porque el nivel de reflexión sobre las tecnologías es todavía limitado y la caja de herramientas de los usuarios es todavía demasiado rudimentaria. Es preciso asumir una actitud crítica y curiosa, desarrollar capacidades sobre todo a nivel individual, comprender de qué formas se puede interactuar con los mundos digitales, poner a punto instrumentos adecuados a los objetivos personales. El desafío consiste en multiplicar los espacios y las ocasiones de autonomía sin ceder ante entusiasmos fáciles, sin sucumbir a la paranoia del control. *Just for fun*. La práctica comunitaria no es una receta capaz de convertir por arte de magia toda novedad tecnológica en un bien colectivo, no basta con conjurar el dominio tecnocrático en nombre de una gran democracia electrónica. Se trata de una visión que confía ciegamente en el progreso, que olvida el valor de las elecciones individuales. La sinergia entre los sujetos en las redes, mundos vivos y constantemente cambiantes, no es una banal suma de las partes en juego, requiere pasión, confianza, creatividad y una continua renegociación de los instrumentos, de los métodos y de los objetivos.

Vincular los elementos más estrictamente técnicos a sus correspondientes efectos sociales es sin duda el primer y el más arduo paso a acometer. Por esta razón, el texto que tenéis en vuestras manos se puede descargar integralmente gracias a una licencia copyleft.

www.ippolita.net
info@ippolita.net

1
Historias de Motores

Motores de búsqueda

Los motores de búsqueda hoy en día se presentan como sitios pertenecientes a la World Wide Web (Web) que permiten hallar informaciones (*informations retrieval*). La mayoría de los usuarios cree erróneamente que Internet y la Web coinciden, porque la Web es el aspecto más sencillo e inmediato de Internet. En realidad la Red es mucho más compleja, heterogénea y variada que la Web: incluye también los mundos de los chats, de los newsgroups, del correo electrónico y, en general, todo lo que los individuos quieren «poner en red», prescindiendo de la «forma» en la que estas informaciones se presenten. Además, la Red no es estática, sino dinámica. Esto significa que las conexiones entre los recursos se modifican, están sujetas a cambios (nacimiento, transformación, muerte). Del mismo modo, los vectores de conexión física que llevan a estos recursos cambian continuamente, evolucionan desde la toma telefónica del módem hasta la «banda ancha» y la fibra óptica. En definitiva, los individuos que crean la Red proyectando sus álter egos digitales en ella están cambiando continuamente, en tanto que seres vivos. La Red, pues, no equivale a la Web, es una dinámica coevolutiva

constituida por las interacciones complejas entre tres tipos distintos de máquinas: máquinas mecánicas (ordenadores, cables, módems, etc.), máquinas biológicas (individuos humanos), máquinas significantes (recursos compartidos).

De todas formas, ya que conseguir información en el magma de la Red ha sido siempre una exigencia fundamental, por mucho que pueda parecer extraño, la historia de los motores de búsqueda es mucho más antigua que la de la Web.

La Web, tal y como la conocemos hoy, es un invento de Tim Bernes-Lee, Robert Caillau[1] y otros científicos europeos y norteamericanos. Entre 1989 y 1991 en el CERN de Ginebra se crean los primeros navegadores, el protocolo HTTP y el lenguaje HTML para la composición y visualización de documentos hipertextuales, es decir, documentos que contienen links (ya sean internos a cada documento ya sean entre documentos distintos). Esta nueva tecnología se presenta como una implementación de la red de Internet, un proyecto estadounidense resultante de la fusión de distintos proyectos universitarios y militares.

Mientras la Web se difundía entre los laboratorios y las universidades de todo el mundo, los motores de búsqueda ya ofrecían, desde hace años, servicios de indexación y de búsqueda de informaciones en Internet.

Los primeros motores de búsqueda obviamente no eran consultables vía Web: eran verdaderos programas instalables y configurables en el propio ordenador. Tales instrumentos indexaban los recursos de Internet mediante protocolos como FTP (para el intercambio de *files*), Gopher (un rival del emergente HTTP) y otros sistemas ya caídos en desuso.

1 Véase Tim Berners-Lee, *Tejiendo la Red, el inventor del Word Wide Web nos descubre su origen*, Siglo XXI, 2000; y *L'architettura del nuovo Web. Dall'inventore della rete il progetto di una comunicazione democratica*, Feltrinelli, Milán, 2001.

En 1994 se hace operativo el primer motor de búsqueda basado en la Web: el WebCrawler[2], un experimento financiado por la Universidad de Washington. Las innovaciones introducidas por este motor de búsqueda eran extraordinarias: además de funcionar como un sitio Internet, y de ofrecer la posibilidad de realizar búsquedas *fulltext*[3], disponía de un instrumento para catalogar las páginas de forma automática, el *spider*. El *spider* es un software que, como veremos en profundidad más adelante, desarrolla dos funciones: mientras navega en Internet, memoriza las informaciones contenidas en las páginas web y las pone a disposición de los usuarios del servicio de búsqueda.

Aunque WebCrawler fuese increíblemente revolucionario para la época, los resultados que era capaz de ofrecer eran simples elencos de direcciones web y los respectivos títulos de las páginas indexadas.

En los últimos meses de 1994 el motor de búsqueda Lycos conseguía indexar en muy pocos meses más del 90% de las páginas de aquel entonces (unos 10 millones aproximadamente). La principal innovación de Lycos era la de excluir sistemas *fulltext* para la catalogación, analizando solo las primeras veinte líneas de toda página indexada, en lugar de examinar el contenido integral del texto de cada página. Esto le permitía a Lycos mostrar en los resultados incluso una pequeña descripción extrapolada de las mismas líneas.

Con la llegada de Excite, en diciembre de 1995, los resultados empiezan a ser ordenados en base al criterio de importancia de las páginas. La introducción de un sistema de «peso» de una página posibilitaba también una forma rudimentaria de catalogación temática: por fin se acababan los interminables listados de resultados sin ordenar. Se configu-

2 Todavía existe y es un óptimo metamotor de búsqueda que combina los resultados de los mejores motores: http://www.webcrawler.com.

3 Gracias a la búsqueda *full-text* es posible encontrar palabras, combinaciones de palabras, frases completas, buscando en el interior de todo el documento.

raba así una primera consulta al *directory* de los sitios, algo comparable a una organización de tipo biblioteconómico clásico de los recursos de la Web, o sea, en forma de árbol, con una indexación dividida por materias, lenguas, etc.

Sin embargo, Excite ha pasado a la historia por otra razón: ha sido el primer motor de búsqueda que se ha equipado de instrumentos dedicados explícitamente a la actividad comercial. Después de haber adquirido WebCrawler, ofreció a los usuarios personalizaciones y cuentas de correo gratuitas, convirtiéndose en menos de dos años en uno de los primeros portales web (1997). En poco tiempo, con el nacimiento de motores de búsqueda cada vez más evolucionados, Excite se aparta de su vocación originaria, optando por el uso de instrumentos de búsqueda de otras sociedades, entre las que hoy figura también Google[4].

El panorama de los precursores de Google no sería completo sin una mención a lo que en 1997 era el mejor motor de búsqueda existente, y también el más popular: AltaVista. AltaVista, que quiere decir «vista desde lo alto», se valía de los resultados de un grupo de investigación de la Digital Equipment Corporation (DEC) de Palo Alto, California, que en la primavera de 1995 había conseguido almacenar todas las palabras de una página HTML presente en Internet en un índice en el que era posible realizar búsquedas. Esto le había permitido el desarrollo de la primera base de datos abierta a la consulta directa de la World Wide Web. La cabeza pensante de AltaVista era Louis Monier[5] que, usando *clusters* de ordenadores conectados entre sí, hardware de vanguardia y lo mejor de las tecnologías entonces disponibles, había conseguido que su creación fuese el motor de búsqueda más utilizado y apreciado por los usuarios. AltaVista ha sido también el primer motor de búsqueda multilingüe en Internet y la primera tecnología de búsqueda capaz de acoger textos en chino, japonés y coreano. Además, ha introducido un sistema de traducción automática, Babel Fish, aún muy utilizado.

Antes de su hundimiento[6] en 1997, AltaVista daba servicio a 25 millones de búsquedas diarias y recogía 50 millones de dólares anuales de los patrocinios, ofreciendo su servicio de búsqueda a los usuarios del portal Yahoo!, que sigue siendo todavía el mayor antagonista de Google en el mundo web.

4 La idea de usar las funcionalidades de búsqueda de otros sitios, y por lo tanto de poder ofrecer entre los resultados propios también los ajenos, no ha sido de Excite. Ya a finales de 1995, MetaCrawler ofrecía en sus páginas los servicios de todos los motores de búsqueda presentes. Hoy en día existen muchísimos metamotores que ejecutan búsquedas cruzadas sobre distintas tipologías de bases de datos.

5 Después, Monier trabajó para Ebay; en 2005 fue contratado por Google.

El nacimiento de Google: en el principio fue el garaje o, más bien, la universidad

La palabra Google procede de «googol», término matemático que indica un 1 seguido por 100 ceros. La leyenda afirma que éste fue el número de páginas web que Larry Page y Sergey Brin soñaban con indexar con su nuevo instrumento de búsqueda. Se conocieron en 1995 en Stanford, cuando Larry Page (tenía entonces 24 años), licenciado en la Universidad de Michigan, visita Stanford con la idea de matricularse en el doctorado de informática. Sergey Brin (entonces 23 años) formaba parte del grupo de estudiantes que tenían que acompañar a los visitantes por el campus. Stanford era el lugar ideal para desarrollar proyectos altamente innovadores desde el punto de vista tecnológico. De hecho podía ofrecer, por un lado, laboratorios de investigación de vanguardia; por

6 En 1996 el valor de esta búsqueda aún no era contabilizado, considerando el estado embrionario en el que se hallaba la economía relacionada con Internet, pero su potencial había sido claramente reconocido por la DEC, sociedad que se encargaba de los asuntos de hardware de AltaVista. Sin embargo, el grupo DEC no será capaz de explotar la ocasión a causa de una serie de desafortunadas reestructuraciones internas que provocan una ruptura en el equipo de AltaVista, adquirido en 2003 por Overture (luego cedida a Yahoo!).

otro, estrechas relaciones tanto con empresas del sector de las tecnologías de la información como con personas atentas y dispuestas a invertir grandes capitales en las investigaciones académicas más prometedoras. Brin y Page están fascinados por el crecimiento vertiginoso de la Web y por los problemas relativos a la búsqueda y a la organización de las informaciones; se dedican al proyecto BackRub, así llamado porque era capaz de analizar los *back links* que remitían a un sitio y de ofrecer un mapa de la situación. Más tarde, BackRub sería rebautizado como Google cuando en 1997 dispusiera de una página web.

La innovación fundamental introducida por Google es la de dar un vuelco al proceso de indexación de las páginas: ya no se enumeran los sitios en base a la «proximidad» respecto a la pregunta realizada, sino que se muestran los resultados exactamente en el orden «correcto», es decir, los que responden a las expectativas del usuario. El primer link es el que da la respuesta «exacta» a la pregunta; los sucesivos se alejan poco a poco del centro de la pregunta[7].

Con esta perspectiva nace el conocido botón «voy a tener suerte»: clicándolo, se abre directamente el primer link de la búsqueda de Google, presentado como el link inequívocamente «exacto».

El algoritmo que Larry Page «inventa» para calcular la importancia de una página web, conocido como PageRank(™), se basa en la matemática estadística de los primeros años del siglo XIX y especialmente en las fórmulas desarrolladas por Andrej Andreevic Markov para calcular, dentro de una red, la importancia de un nodo respecto a los demás[8].

7 Creemos que este enfoque es absolutamente irreal, además de ser políticamente incorrecto e intrínsecamente autoritario, ya que supone la existencia de una autoridad capaz de individualizar y ofrecer una respuesta «correcta, rápida y justa» a un deseo. Pueden existir muchos recorridos individuales para construir una respuesta aceptable y satisfactoria a una necesidad/deseo (véase capítulo V).

8 Véase capítulo IV.

En un primer momento, Google es únicamente un proyecto académico y científico, para el que el sistema de evaluación depende en gran medida de los juicios de los reseñadores (*referees)* que operan según el mecanismo de la «revisión por pares» *(peer-review)*. El método que teóricamente ofrece las mayores garantías de objetividad es el llamado «doble ciego» *(double-blind)*: antes de ser publicado en una revista científica, un artículo se somete a la evaluación de por lo menos dos investigadores de renombre, que no deben conocer el nombre del autor (para no verse influenciados en su juicio). El segundo paso «ciego» tiene que ver con el redactor del artículo, que no debe conocer el nombre de los reseñadores. En suma, cuanto más positiva es la acogida que un artículo recibe por parte de otros científicos (que se supone trabajan de forma independiente), tanto más el artículo se considera importante y meritorio. Page traslada a su campo de búsqueda este enfoque y desarrolla la teoría según la que el número de links que conducen a una página es una forma de evaluar el valor de la página y, en cierto sentido, su calidad. Más adelante veremos mejor cómo se realiza este paso de la «cantidad» de informaciones suministradas a la «calidad» del resultado que da respuestas a las expectativas del usuario[9].

Sin embargo, este criterio no era de por sí suficiente porque no todos los links son iguales ni tienen el mismo valor: más exactamente, desplazando la atención del valor estático de una conexión al valor dinámico de un recorrido, la Red es un ambiente (en términos matemáticos un *grafo*) en el que no todos los recorridos tienen el mismo valor, sino distintos «valores de rutas» según el «peso» de los nudos. Por otra parte, siguiendo con el paralelismo con la práctica científica, no todas las reseñas en las revistas científicas tienen la misma importancia. La opinión positiva de reseñadores no prestigiosos, o incluso mal vistos por la comunidad científica, es nociva: demasiadas reseñas de escaso peso invalidan la auto-

9 Véase capítulo VI.

ridad de una publicación. Así, según Page, los links que proceden de sitios especialmente conectados, con un elevado número de links, son de mayor importancia. Por lo tanto, un recorrido (link) que lleva a una página desde un sitio muy popular tiene mayor relevancia que un recorrido procedente de una página poco conocida. De esta manera, un link desde una página A a una página B se interpreta como una reseña científica (aunque hay que subrayar que Brin y Page hablan explícitamente de «voto» y «democracia»), de mayor peso dependiendo de la popularidad del reseñador/votante que facilita el link. El criterio de la autoridad del reseñador científico se convierte banalmente en la medida de la popularidad de un sitio.

La evaluación de las páginas web de Google, llamada *pageranking*, se construye pues en base a un sistema de reseña «pública» análogo al sistema de la «república científica»[10], que no está limitado a los científicos sino que se extiende a todos los internautas de la Red.

Hoy en día, la organización de la república científica y la cuestión de la localización de las informaciones en general se han convertido en un problema de importancia crucial: en este contexto de diluvio informacional[11], particularmente en

10 El sistema de las publicaciones científicas nace en 1665 con *The Philosophical Transactions*, la primera revista científica fundada por Henry Oldenburg de la Royal Society de Londres, una especie de registro público de la propiedad intelectual; funcionaba como una oficina de patentes de las ideas científicas. La revista quería introducir claridad y transparencia en el proceso de fundación de las pretensiones innovadoras y tales pretensiones se veían consolidadas gracias al reconocimiento de un título de «nobleza intelectual» por parte de unos pares, lo cual garantizaba un título especial de propiedad: la intelectual. La importancia de las publicaciones contribuía a la creación de una «jerarquía entre pares», basada en un ideal de excelencia: sin embargo, ya que los resultados para ser apreciados necesitaban publicidad, la organización de las reglas de publicación asumió un papel central, ya que para un científico ser visible se convirtió en algo fundamental para adquirir notoriedad y prestigio.

11 El concepto de «diluvio informacional» (*déluge informationnel)* ha sido desarrollado sobre todo por Pierre Lévy, reelaborando la idea de «segundo diluvio» del artista Roy Ascott. De ahí procede un cuadro filosófico en el que el movimiento de «virtualización» asume un papel central, véase: Pierre Lévy, *¿Qué es lo virtual?*, 1998, Ediciones Paidós Ibérica, S.A. (online la versión original francesa: http://hypermedia.univ-paris8.fr/pierre/virtuel/virt0.htm). De manera independiente, Manuel Castells habla de «informacionalismo» y de «economía informacional». Gran parte de las lecturas económicas acerca de la época del diluvio informacional, especialmente en el ámbito francés, italiano y estadounidense, tienen derivaciones marxistas y descansan sobre interpretaciones forzosas de los fenómenos de «inteligencia colectiva» de las redes como hipóstasis del *General Intellect* marxista; véase por ejemplo: Olivier Blondeau, *Genèse et subversion du capitalisme informationnel*, http://severino.free.fr/archives/copieslocales/geneseetsubversion.html; la revista monográfica *La società del General Intellect*, Derive approdi, Roma; Wark McKenzie, *Un manifiesto hacker*, 2006, Ediciones Alpha Decay S.A.

la Web, razonar no solamente sobre la importancia sino también sobre la credibilidad de la información es cada vez más complejo. Y esto lo es mucho más desde que el mecanismo de la «revisión por pares» ha sido puesto en tela de juicio por los propios científicos[12]. Entre las propuestas alternativas más interesantes figuran mecanismos de evaluación distribuida, puestos a disposición gracias a licencias libres (*copyleft*) me-

12 Sitiografía sobre la revisión por pares. Definiciones: http://es.wikipedia.org/wiki/Revisi%C3%B3n_por_pares; la topología de una red de citas: Stefano Mazzocchi, «Peer Review vs. Citation Network Topology», http://www.betaversion.org/~stefano/linotype/news/74/; artículos críticos recientes, *The scientist*: Alison McCook, «Is Peer Review Broken?», *The Scientist*, vol. 20, tomo 2, p. 26, http://www.the-scientist.com/2006/2/1/26/1/; Richard Gallagher, «Taking on peer review», *The Scientist*, vol. 20, tomo 2, p. 13, http://www.the-scientist.com/2006/2/1/13/1/; Google y la revisión por pares: http://www.lib.uiowa.edu/hardin/md/notes4.html; el fenómeno de los blogs entre informaciones y rumores de fondo: Fabio Giglietto, http://nextmedia.blogspot.com/2004/10/informazionenon-informazione-fra.html; citas y revisión por pares, ciencias duras y ciencias suaves: Francesca di Donato, «Verso uno European Citation Index for the Humanities – Che cosa possono fare i ricercatori per la comunicazione scientifica», http://bfp.sp.unipi.it/~didonato/ovre/ECIH.html; el exhaustivo trabajo de Jean-Claude Guédon, «In Oldenburg's Long Shadow: Librarians, Research Scientists, Publishers, and the Control of Scientific Publishing», Univerisité de Montreal, http://www.arl.org/arl/proceedings/138/guedon.html.

diante proyectos de acceso público (*open access*), incluso por lo que respecta a obras de humanidades (véase por ejemplo *HyperJournal*)[13].

Así en marzo de 1996 Page lanza su spider, un programa para explorar la Web, testándolo con el algoritmo de ranking que había desarrollado.

El motor de búsqueda de los dos talentos de Stanford, construido alrededor del spider, se convierte inmediatamente en un éxito entre los estudiantes y los investigadores, adquiriendo rápidamente una extraordinaria popularidad. Sin embargo, para los administradores de la Red de Stanford el consumo de banda del motor de búsqueda empieza a convertirse en un problema significativo. Además, los propietarios de los sitios indexados se preocupaban, por un lado, de los derechos de copyright de sus propios materiales y, por otro, del hecho de que el ranking de Google no tomara en consideración alguna los premios oficiales o los demás sistemas de evaluación consolidados, salvo por la cantidad y calidad (en el sentido de popularidad) de las relaciones que una página es capaz de mover a su alrededor. A Google le interesa la economía relacional expresada en términos de link y nada más: «el spider no se interesa por los contenidos de una página».

Una afirmación de este tipo supone que el valor de un resultado se construye mediante la evaluación de las relaciones de una página con otras, y no imponiendo una clasificación

13 Sitiografía: sobre el acceso a los saberes, la libertad de información y el acceso abierto a la literatura científica. Destacamos el libro del Gruppo Laser, *Il sapere liberato*, Feltrinelli, Milán, 2005, online en www.ippolita.net/laser/. En el campo humanista, el trabajo desarrollado por Francesca di Donato, que consigue conjugar un enfoque cualitativo y no cientifista con la necesidad de descentrar los archivos y las informaciones, y definir un número mínimo de criterios a priori de catalogación (metadatos, por ejemplo XML), para intercambiar mediante un protocolo común. Véase «HyperJournal. Un software open source per l'open publishing» (abstract y slides), http://www.hjournal.org/taxonomy/term/2; seminario: «Dopo Berlin 3: politiche di accesso aperto alla letteratura scientifica» (Pisa, 16 de febrero 2006), http://eprints.rclis.org/7650/2/borellislight.pdf. El sitio de *HyperJournal*: http://www.hjournal.org/.

arbitraria en base a términos de búsqueda. Este forzamiento será la clave de la eficiencia de Google en los siguientes años: los resultados no se establecen de manera absoluta una vez por todas, sino que varían dinámicamente según la posición que asumen las páginas en el contexto de la Red.

Google.com: publicidad directa dentro de las páginas

Page e Brin desarrollan y prueban Google durante todo el primer año y medio de vida usando instrumentos libres puestos a disposición por las comunidades free software y open source[14], a partir del sistema operativo GNU/Linux. De esa manera implementan un instrumento fuertemente modular y escalable, que se puede expandir y modificar fácilmente incluso mientras está siendo usado. Esta estructura modular hoy en día es la base del *datacenter*[15] de Google, el Googleplex, que permite sustituir o añadir hardware y software, además de realizar su manutención y actualización sin tener que bloquear nunca el servicio.

A mediados de 1998, Google da servicio a cerca de diez mil búsquedas diarias; el sistema casero apañado en la habitación alquilada por Page y Brin se halla al borde del colapso. Era preciso, pues, encontrar un sistema de financiación para el proyecto que fuese más allá de los normales fondos de ayuda para la investigación.

14 Véase en el capítulo III el análisis más detallado de las relaciones entre estas fórmulas.

15 El datacenter de Google se compone actualmente de varios datacenters esparcidos por el mundo que elaboran las solicitudes realizadas y se encargan de propagar las modificaciones de las indexaciones. La actualización de las bases de datos, así como el cálculo del pageranking se resiente de esta descentralización: los datos normalmente se alinean en unos días, o algo más; este desfase da lugar a la llamada «Google dance», el baile de los ranking. Para una panorámica general, véase el artículo http://www.alverde.net/articoli/stampa.asp?ID=145 y http://www.motoricerca.info/articoli/google-dance.phtml sobre la «Google dance».

La salida de la universidad se atribuye al encuentro con Andy Bechtolsheim, fundador de la Sun y consumado *talent scout* de las tecnologías de la información, que entrega a los dos investigadores un primer cheque de cien mil dólares.

El nacimiento de la sociedad coincide con la contratación de los primeros empleados necesarios para el desarrollo y el mantenimiento del datacenter, entre los que figura Craig Silverstein, director de tecnología. El datacenter de Google se configura desde el principio como un sistema altamente redundante, en el que las informaciones se copian en más lugares para minimizar la posibilidad de pérdida de los datos (o sea, el papel moneda de un sistema de búsqueda), y sobre todo un sistema en el que en todo momento es posible añadir o quitar módulos existentes para conseguir una mayor eficacia. El enfoque típico de la filosofía de los dos lumbreras universitarios, Brin y Page, que recuperan y readaptan hardware barato y utilizan software gratuito, se revela una carta ganadora. Los reducidos recursos económicos permiten luego convertir este sistema en la piedra angular de su filosofía empresarial, que se basará en la extrema escalabilidad a todos los niveles. El sistema Google es *escalable* porque es capaz de «crecer» o «decrecer» (aumentar o disminuir de escala) conforme a las necesidades y a las disponibilidades. No es necesario replanear el sistema cuando se añaden nuevos recursos (de cualquier tipo: máquinas, servicios, empleados, etc.): la estructura altamente modular permite la adquisición de nuevos módulos sin solución de continuidad.

Google Inc. abre sus puertas el 7 de septiembre de 1998 en Menlo Park, California. Más exactamente, Larry abre estas puertas con un mando a distancia: las oficinas se hallan en el garaje que un viejo amigo subalquila a la sociedad. Una oficina-garaje espartana, pero dotada de algunos lujos no tan insignificantes: una lavadora, una secadora e hidromasaje. Desde el principio, la filosofía empresarial de Google se basa en la satisfacción de los trabajadores.

En enero de 1999 Google abandona definitivamente el campus de Stanford, declarando: «El proyecto de investigación Google se ha convertido en Google Inc. Queremos darle al mundo búsquedas de calidad superior y mejores de las que hay actualmente, y una compañía parece ser la mejor forma para llevar a cabo este objetivo. Hemos empezado a incorporar a otras personas y a configurar otros servidores para hacer que el nuestro sea un sistema escalable (hemos empezado a ordenar los servers en paquetes de 21). Hemos empezado también a lanzar nuestro spider con más frecuencia y nuestros resultados no solo siguen siendo tan rápidos como antes, sino que están aún más actualizados. Estamos incorporando personas de gran talento muy rápidamente y esto comporta disponer de las últimas y las mejores tecnologías por lo que atañe a la Web»... Brin y Page dedican luego algunas líneas a las diez buenas razones para trabajar en Google, incluyendo tecnologías, acciones, *snacks*, bebidas gratuitas y la posibilidad de que millones de personas «utilicen y aprecien vuestros softwares».

Entre 1998 y 1999 todos los motores de búsqueda y los sitios más populares del mundo están afectados por el síndrome del portal, una verdadera manía en la construcción de sitios que quieren que los visitadores se queden a cualquier precio más tiempo en la propia página, para ello incluyen cada vez más servicios, publicidad y personalizaciones. Google es el único instrumento carente de publicidad y de servicios añadidos: solo motor de búsqueda, entre otras cosas el mejor, el más rápido y sin lazos comerciales o patrocinadores.

En todo caso, la compañía no puede sobrevivir con la modesta suma de Bechtolsheim, sin generar aún ganancias significativas y siguiendo con la investigación sobre la organización y la localización de las informaciones. Mostrando una gran capacidad de interlocución con las altas esferas de las finanzas e insistiendo en su interés por la investigación, Brin y Page consiguen poner de acuerdo a las dos principales sociedades financieras de California, rivales entre sí, pero que sorprendentemente aceptan financiar conjuntamente a Goo-

gle, caso tan único como raro: dos enormes sociedades de capital de riesgo se dejan convencer para compartir en lugar de competir por los riesgos y las ganancias del eventual negocio. El 7 de junio de 1999 Google puede anunciar que la Sequoia Capital y la Kleiner Perkins Caufield & Byers le han concedido una financiación de 25 millones de dólares[16].

Mientras van aumentando los empleados de la Google Inc., los dos presidentes buscan una forma para explotar económicamente la gran cantidad de datos indexados. Inicialmente intentan vender su servicio de búsqueda como OEM (*Original Equipment Manufacturer* – productos ofrecidos sin los soportes originales)[17] a portales, pero los resultados son más bien escasos. El modelo de negocio que se revela más adecuado para la nueva sociedad es la publicidad directa dentro del motor de búsqueda, mediante el cálculo del número de visitantes que acceden a los sitios desde los links promocionales y de pago. Este modelo económico, CPM (coste por mil impresiones; en inglés CPT, *cost per thousand impresions*)[18] se estructura de manera tal que sea lo menos invasiva posible para el usuario; de hecho, en lugar de optar por la extrema visibilidad de los anuncios publicitarios, se muestran solo unos pocos links cuidadosamente seleccionados en la parte de arriba de la búsqueda. Estos links tienen colores y caracteres distintos de los que no son comerciales, pero intentan no molestar de ninguna forma la actividad de búsqueda del usuario.

16 Ambas sociedades detentan aún hoy las cuotas mayoritarias del capital de Google.

17 Una OEM es una sociedad que construye componentes utilizados en sistemas vendidos por otra sociedad, por ejemplo, un revendedor. Sociedades que venden productos con componentes que proceden de OEM son comunes en la industria electrónica e informática. Por lo general, la OEM se contrapone al RETAIL porque el producto se ofrece a terceros sin los soportes originales (CD de instalación, manuales, etc.), integrado en el producto por el revendedor.

18 CPM, CPC, etc.: los modelos de negocio publicitario online son cada vez más complejos y refinados. Para una rápida introducción véase http://earlystagevc.typepad.com/earlystagevc/2005/08/web_20_needs_bu.html.

El autoservicio publicitario: más allá de la burbuja de la *new economy*

Un modelo económico basado en la aparición de simples links a los que patrocinar, colocados al lado de las búsquedas, es poco ventajoso en términos de ganancias; por lo que a esto se refiere, la planificación empresarial a largo plazo de Google necesita de un salto de calidad. Los presidentes miran a su alrededor en busca de las soluciones más interesantes y beneficiosas, y se topan con Goto, sociedad fundada por Bill Gross[19], hoy propiedad de Overture/Yahoo!.

El negocio de Goto consistía en mezclar resultados efectivos y resultados patrocinados, facturando a los anunciantes solamente los clics realizados por los usuarios en sus direcciones: un modelo conocido como CPC (*cost per clic*).

Este método era especialmente innovador respecto a los anteriores. Los links patrocinados aparecían a continuación de las búsquedas de los usuarios, ampliando de esa forma al máximo las posibilidades de llevar a buen término la transacción económica, contabilizada como clic en el sitio pagador. Google intenta asociarse con Goto, pero el decidido rechazo del presidente les obliga a implementar internamente un instrumento parecido. En esa época todos los portales (por ejemplo Excite, Lycos, Infoseek, AltaVista y Yahoo!) utilizaban el modelo CPM, y el CPC era casi un deseo prohibido. Esto demuestra como un intento fracasado de adquirir la tec-

19 La historia de Goto es muy emblemática porque muestra como incluso en el ámbito de la Web cada vez más se va hacia las megafusiones típicas del capitalismo informacional. Un buen resumen en italiano hasta el 2004 se halla en http://motoridiricerca.it/goto.htm; en wikipedia, recursos sobre el papel comercial jugado por Goto/Overture en el negocio de Yahoo! en http://en.wikipedia.org/wiki/Yahoo!_Search_Marketing.

nología que se cree mejor para el propio negocio conlleva la necesidad de implementarla de forma autónoma.

En marzo del 2000 estalla la burbuja del Nasdaq, arrastrando consigo todos los sueños de gloria de las dot.com. También se viene abajo el modelo CPM, es decir, la ilusión de que toneladas de banners publicitarios a «miles de impresiones», muy a menudo descontextualizados por estar presentes en sitios totalmente ajenos al negocio de referencia del anunciante, fueran a ser una fuente inagotable de ingresos. Google tiene entonces que enfrentarse con la necesidad de monetarizar sus propias tecnologías de búsqueda.

En octubre del 2000 entra en escena AdWords, una especie de autoservicio publicitario en el que los anunciantes pueden elegir las palabras de búsqueda a las que asociar su propio link comercial: una implementación de la «publicidad por palabras claves» (*keywords advertising*) ideada por Goto.

Google sobrevive por lo tanto al hundimiento del Nasdaq e incluso, gracias a su naturaleza de sociedad privada que aún no cotiza en bolsa, aprovecha para contratar todos los talentos en fuga de las ruinas de las demás dot.com. A mediados de 2000, Google da servicio a 18 millones de búsquedas diarias y su índice de documentos cuenta con mil millones de unidades: seis meses después las búsquedas ascienden a 60 millones.

Estilo, morfología y multiplicación de los servicios

A principios del 2000 Google tiene muy pocos competidores y da comienzo a una nueva fase de innovación, empezando a ofrecer una serie de nuevos servicios a los usuarios[20].

Cada nuevo servicio forma parte de un mosaico complejo, constantemente redefinido, con ramificaciones en todos los

20 Aquí aludiremos solo a los servicios más interesantes; para un listado exhaustivo de los servicios ofrecidos por Goolge, véase http://www.google.com/intl/en/options/ (los servicios disponibles en castellano y en otros idiomas son menos, pero en rápido aumento).

ámbitos de las tecnologías de la información. Actualmente Google ofrece: diecisiete tipologías de servicios de búsqueda en bases de datos de imágenes, blogs, noticias, libros, mapas, vídeos, servicios financieros, servicios de búsqueda en los documentos presentes en el propio ordenador y mucho más. Dos servicios están relacionados con el desarrollo de aplicaciones y con las novedades sobre las que se está trabajando en los laboratorios de Google; y seis con los servicios de comunicaciones: e-mail, VoIP, mensajería instantánea, grupos de debate, fotos compartidas, servicios de traducción. Tres de los servicios tienen que ver con dispositivos móviles (móviles, PDA, etc.); y, para terminar, un servicio propone softwares aconsejados por Google. No hay día en que no se añadan nuevos servicios.

Hasta el usuario más ignorante se da cuenta fácilmente del poder de estos instrumentos. Hoy es posible digitar una dirección o un número de teléfono, y Google ofrece todos los datos necesarios para contactar y localizar a una persona o una cosa; se pueden memorizar las preferencias personales de búsqueda de modo tal que la búsqueda sea de una facilidad desconcertante; si hay un error en la digitación de la palabra en la búsqueda, Google sugiere la búsqueda correcta con un mecanismo de corrección ortográfica muy avanzado capaz de «aprender» de cada búsqueda.

En 2001 Google lanza «Google imágenes», un motor de búsqueda de imágenes que en pocas semanas se convierte en una de las principales fuentes de producción gráfica del *do it yourself* y representa una de las mayores base de datos de imágenes disponible en la Web. Mientras tanto la compañía compra a Deja.com el archivo Usenet, es decir, los mensajes de los newsgroups desde 1995, con más de 650 millones de *posts* es la «memoria histórica» de Internet antes de la Web, cuando los grupos de debate eran la linfa vital de la Red. En abril de 2001 Usenet se convierte en «Google grupos», con una interfaz bastante simple y cautivadora para seguir con facilidad complejos temas de debate. Todos estos servicios están en continua ampliación y mejora.

Desde 2001 en adelante los avances y servicios se multiplican sin ningún objetivo evidente ni contraprestación económica inmediata, como si Brin y Page se divirtieran mostrando que un datacenter ilimitado es capaz de elaborar cualquier sueño tecnológico. El ejemplo más desconcertante es el poderosísimo Google Maps (Tierra, Marte 2005 y Luna 2006), un set de software disponible gratuitamente para visualizar con imágenes de satélite un mapa detallado o cuando menos una reconstrucción fotográfica de la superficie de los planetas. Google Directory presenta los contenidos *open directory* de Dmoz.com, organizados por agentes humanos con un sistema cooperativo y descentralizado de *open publishing*, y desde hace tiempo consultable en la gráficamente cada vez menos lacónica página de Google.

En 2005 aparecen Google News, la enorme base de datos de informaciones de Google que se pone así al servicio de la producción periodística, y Gmail que ofrece 1 gigabyte de espacio de mail a cada usuario. Con un mecanismo de invitaciones que crea inmediatamente una red de relaciones interna al mundo de Google, se silencian las protestas acerca de la privacidad de la propia cuenta electrónica con una frase lapidaria: «se trata de un buen producto, las ventajas son mayores que las dudas y el tiempo lo demostrará». Nadie de la competencia tiene la capacidad de ofrecer gratuitamente tanto espacio. Con Gmail cada actor implicado está controlado desde el punto de vista del uso del servicio, ya que la enorme disponibilidad de espacio invita a dejar los correos propios en los servidores de Google. Además, puesto que la difusión del servicio se da mediante invitaciones, que cada usuario puede usar a su gusto, a Google se le facilitan una serie de informaciones esenciales sobre los conocidos y los amigos: un proyecto invasivo y enfocado al *data mining*, el monitoreo de los datos.

Después llega el proyecto Google Scholar (Academia), un motor de búsqueda todavía experimental en fase beta (la llamada fase de *beta-testing*) que permite localizar textos de la literatura académica así como artículos a revisar, tesinas, tesis doctórales, volúmenes académicos, *preprint*, sumarios, bibliografías, informes, reseñas de todos los sectores de la investigación académica. Y, finalmente, Google Library, que ambiciona almacenar online todos los libros del mundo en formato digital, sellando acuerdos con bibliotecas de todo el mundo para la digitalización de los volúmenes, pero también con las editoriales que estén interesadas. Un sueño que solamente el datacenter de Google promete poder realizar: la biblioteca digital global accesible desde las páginas de Google, en abierta ruptura con la mayoría de las editoriales estadounidenses, reunidas en la Association of American Publishers, que temen una drástica caída de las ganancias; ya en 2005 la AAP había solicitado el bloqueo durante seis meses de la digitalización de las obras protegidas por copyright a la espera de explicaciones convincentes sobre la Google Library. Pero las apariencias engañan y, a pesar de la oposición de los partidarios del copyright, no se trata de una iniciativa para la libre circulación de los saberes, sino solo de un cambio de monopolio de la información, que pasaría de las grandes editoriales a Google. De hecho, como pasa con todo sueño, la pesadilla está al acecho tras la esquina: un sujeto privado, Google, ambiciona decidir qué va a formar parte de la esfera de la información colectiva, poniéndolo a disposición mediante formatos propietarios. Como reacción a este proyecto ha surgido la Open Content Alliance, apoyada por Internet Archive, una fundación sin ánimo de lucro, y por Yahoo!: el objetivo es la disponibilidad total del material adquirido mediante el uso de formatos abiertos.

Paralelamente al nacimiento de nuevos servicios, Google se revela extremadamente hábil en aprovecharse plenamente de la economía relacional, generada por la explotación en sentido comercial de los datos indexados.

En 2004 se lanza AdSense, es decir, la posibilidad para cualquiera de hospedar en su propio sitio algunos links comerciales sugeridos por Google basados en los temas tratados y en algunas palabras claves. Los ingresos procedentes

de los clics en estos links se reparten entre Google y los propietarios de los sitios hospedantes. La novedad de este enfoque consiste en explotar la red de confianza basada en los usuarios del sitio: Google ya no está en el sitio de Google, sino que se halla en cualquier sitio que quiera hospedar su «ventana», y este pequeño espacio ofrecerá siempre datos válidos e interesantes, como pasa siempre con Google, aunque ahora se tratará de sugerencias comerciales. En la práctica AdSense sustancia una «red de Google», una red que cruza los datos de los usuarios y las relaciones entre ellos para las sugerencias publicitarias. En palabras de Google, es una red que «está formada por un conjunto amplio de sitios web y otros productos, como programas de correo electrónico y blogs, que se han asociado a Google para publicar anuncios de AdWords»[21]; ¡naturalmente el sistema AdWords también forma parte de la «red de Google»!

Desde luego, una vez creada la red, es preciso sacarle dinero. No es casualidad que, aún en el 2005, Google lleve a cabo un «regreso» al modelo CPM a través de la oferta de AdWords con la metodología «targeting por sitio»: los anunciantes vuelven a pagar por visualizaciones (*impresions*) con un mecanismo de subasta y no por número de clics en el banner. Pueden elegir de modo extraordinariamente detallado qué características tienen que reunir los usuarios que visualizarán el anuncio: nacionalidad, área geográfica de procedencia, idioma, etc. Pero, sobre todo, las visualizaciones solo se dan dentro de la «red de Google». Esta estrategia les interesa sobre todo a quienes desean visibilidad y promoción de una firma, más que llegar a usuarios a los que vender un producto, o sea que privilegian una forma indirecta de marketing frente a formas directas. Más en detalle, se trata de invertir en la llamada «conciencia de marca» (*brand awareness*) en lugar de en productos específicos indicados por palabras

21 Definición de la red de Google según Google: http://adwords.google.es/support/bin/answer.py?answer=6104&ctx=sibling.

claves (como, por ejemplo, el *keyword advertising* del modelo CPC).

Todo está rodeado de un sórdido carácter especular entre el modo de gestionar los propios productos inmateriales y la organización de los empleados y del ambiente de desarrollo de los proyectos, perfectamente adecuado al criterio de escalabilidad modular sobre el que se fundamenta la filosofía empresarial de Google: el crecimiento infinito es preciso para que el sistema no se venga abajo. El número de usuarios que realizan búsquedas en Google y confían a sus servidores sus propios datos tiene que crecer continuamente; para que crezca a la par el número de anunciantes que publican en la «red de Google» sus propios productos son necesarios siempre nuevos servicios, nuevas máquinas para gestionarlos, nuevos empleados para implementarlos, nuevos usuarios para utilizarlos, nuevos anunciantes para generar beneficios... Cada nueva «pieza» del sistema será introducida como nuevo módulo, en un ciclo de acumulación sin fin: acumulación de datos y acumulación de cerebros y acumulación de usuarios y de sus datos, bondad en el tratamiento de los datos, bondad en el tratamiento de los empleados, bondad en el tratamiento de los usuarios y de sus datos, archivados en el datacenter de Google. Siempre con el imperativo de la rapidez y del desarrollo.

Brin y Page no ocultan sus ambiciones: «¿Para qué dejar que nuestros empleados intenten fundar su propia sociedad para luego tener que comprarla si en lugar de eso podemos financiarlos para que se queden con nosotros haciendo lo que habrían hecho de todas formas?». De hecho, el GooglePlex[22], la sede operativa para los trabajadores de Google en Mountain View, es una especie de campus universitario en el que las personas reciben constantes mimos y en el que a los empleados se les deja incluso un día a la semana para que se dediquen a trabajar en sus propios proyectos; estos proyectos

22 Véase el servicio fotográfico del *Time* sobre Googleplex: http://www.time.com/time/photoessays/2006/inside_google/.

se presentan luego al «Google Duo», el cual premia con dinero y apoyo empresarial a los talentos más prometedores.

Google, el gigante bueno, cotiza en Bolsa

Don't be evil, no seas malo, puedes hacer cualquier cosa, mientras no hagas nada malo. Éste es lema del capitalismo humano de Google[23]. A pesar de estas manifestaciones de buena fe, ya se van registrando muchas excepciones en esta campaña a favor del bien: juicios de diversa naturaleza, sospechas de fraude, ocultación de sitios web, etc[24].

En diciembre de 2002 Google cuenta con más de 1.000 empleados y posee más de 10.000 ordenadores. Indexa más de cuarenta mil millones de documentos y las ganancias netas (aunque reveladas con mucha cautela) rondan los 185 millones de dólares. Dadas las dimensiones, los inversores empiezan a exigir mayor claridad y control, imponiendo una estructura societaria más creíble, en cuyo vértice se sitúan los dos geniales y excéntricos ingenieros, pero también un mánager de comprobadas capacidades empresariales. Después de una serie de contactos infructuosos y desencuentros verbales, se asigna el rol de consejero delegado (CEO, *chief executive officier*) de Google a Eric Schmidt (con anterioridad en la plantilla de Sun y luego de Novell).

Los dos jóvenes fenómenos continúan tomando decisiones en todos los campos, aunque la acción diplomática se revela una buena elección económica, puesto que la llegada de

23 La filosofía de Google Inc.: http://investor.google.com/conduct.html.

24 Tres ejemplos de la cantidad de informaciones al respecto: un artículo de WiRed, «Google vs. Evil», http://www.wired.com/wired/archive/11.01/google_pr.html; un simpático cuento de Paul Ford, *The banality of Google* (en lugar de la notoria banalidad del mal...), http://ftrain.com/GoogleIP.html, y un portal entero dedicado al lema, aplicado a un auspiciado retorno de la confianza en los negocios, en la política, en los medias, con declaraciones para que Google permanezca fiel a sus principios, http://www.dontbeevil.com/.

Schmidt coincide con el primer trimestre en activo de la compañía, la cual demuestra que por fin ha cruzado el umbral de la monetarización de sus productos.

Page y Brin habían retrasado al máximo la necesidad de cotizar en Bolsa, temiendo que la obligación de dar cuenta pública de sus perspectivas y ganancias haría su trabajo mucho más difícil, poniendo entre otras cosas a Google en situación de poder ser mejor y más fácilmente controlable y, por lo tanto, vulnerable a sus competidores en el mercado.

Pero ahora, después de AdSense, en 2004, a pesar de que en palabras de Page «Google no es una compañía convencional. Y no queremos serlo de ninguna de las maneras», por razones legales el nuevo coloso se vuelve exactamente como todas las otras sociedades norteamericanas y cotiza en Bolsa.

Precisamente tras la oferta pública de acciones, Yahoo! y otros competidores emprenden decenas y decenas de juicios por violación de patente, de copyright y mucho más contra Google para dañar el nombre ante la perspectiva de la venta pública de las acciones.

Wall Street está a punto de bajar el precio mínimo previsto para las acciones, considerada la dificultad de la operación, pero Page y Brin resuelven la contienda principal con Yahoo!, pagando al coloso de Filo y Yang una compensación en acciones y arreglando las controversias sobre patentes. Los dos, a pesar del parecer contrario de la Bolsa, siguen con la oferta pública de las acciones, justo en la mitad de agosto de 2004, aunque con una rebaja del precio de más de 20 dólares por acción.

En un solo día de contrataciones en el mercado las acciones de Google pasan desde la base de 85 dólares a 100 dólares, con una ganancia neta de 1.000 millones de dólares. Un año después las acciones alcanzan el valor de 400 dólares, con una subida de más del 300%. Para Google Inc. todo parece marchar viento en popa, en el maravilloso mundo donde nadie es malo, todos ganan, nadie se hace daño nunca. Por supuesto, frente a tales cifras, la más mínima oscilación hacia

abajo quiere decir millones de dólares convertidos en humo, tal y como se vio en marzo de 2006 con una inflexión de siete puntos. Google Inc. ya es un gigante entre los gigantes de la Bolsa mundial, y si tiembla, muchos tiemblan con él.

Google Inc.: el monopolio de la investigación

En otoño de 2004 Page y Brin, de viaje en su chárter privado, se enteran de que el coloso AOL (America Online, el mayor proveedor estadounidense) ha firmado un contrato para su propio motor de búsqueda con Yahoo!. Los jóvenes treintañeros cambian de rumbo, aterrizan en Londres y convencen a AOL de que rompa el contrato que acaba de firmar y se pase a Google con una oferta irresistible de 50 millones de dólares. No se trata exactamente del comportamiento abierto y amable que se podría esperar del gigante bueno intentando derrotar los malos monopolistas, pero, sea como sea, los negocios son los negocios, también para los dos valiosos investigadores de Mountain View.

Mientras tanto las ganancias de Google Inc., en un periodo de cinco años, han crecido más del 400.000%, entrando en competencia directa con Yahoo! y Microsoft, no solo en términos de movimiento financiero, sino sobre todo por la popularidad y en consecuencia por el dominio cultural sobre el imaginario colectivo. Millones de usuarios de la Red utilizan las páginas de Google como punto de partida para sus navegaciones, confían en los resultados obtenidos con las herramientas perfeccionadas en Mountain View. Google ya es sinónimo de Web, cuando no incluso de Internet. Google puede aprovechar al máximo la red relacional de sus usuarios y consigue exprimir cada céntimo de millones de pequeños inversionistas, hasta el punto de que en 2005 los datos hablan de cerca de seis mil millones de dólares de ingresos publicitarios (Yahoo! está parado en la cuota de cuatro mil seiscientos millones de dólares).

El primer *impasse* al que se enfrenta Google es la relación entre sus búsquedas y la legislación acerca de las marcas registradas (*trademark*,™). Resultan elocuentes las denuncias interpuestas por Geico y por la American Blind & Wallpaper Factory[25]. En los dos casos las sociedades acusan al servicio AdWords del motor de búsqueda de vender términos bajo su *trademark* a su competencia, violando las leyes sobre marcas registradas. La cuestión es peliaguda, pues las empresas piden a Google que no aparezcan anuncios de su competencia, sino solo los expresamente autorizados, cuando los usuarios buscan términos como «geico», «american blind» o «american wallpaper». Si esta cuestión llegara a los tribunales, una de las más importantes fuentes de ingresos de Google y de sus socios peligraría enormemente, porque permitiría a cualquiera que posea una marca registrada vetar su uso para AdWords o denunciar a Google por su utilización. Luis Vuitton en Francia lleva adelante una demanda similar y gana. La respuesta oficial de Google es que los responsables de la violación, a lo sumo, son los anunciantes y, por supuesto, no Google —que en este caso es un mediador neutro—; y que «limitar la venta de términos registrados significa una limitación de la libertad de expresión». Una posición fuerte de parte de la sociedad y aparentemente más que compartible.

En todo caso, la propia sociedad de Mountain View cae en la misma trampa tendida a las sociedades sobre la libertad de expresión, violando el vínculo de confianza que une a muchos usuarios y que constituye uno de sus recursos fundamentales. Google se ha atrincherado siempre detrás de la supuesta objetividad de los algoritmos de búsqueda y de la objetividad de las máquinas que proporcionan resultados en los que no se

25 De momento solo Geico ha perdido, en los demás casos Google ha sido condenado a una multa. Apuntamos aquí algunas fuentes en italiano acerca de estas vicisitudes judiciales: el caso American Blind and Wallpaper Factory: http://punto-informatico.it/p.asp?i=46751; Google condenado a indemnizar a Luis Vuittón en Francia: http://webnews.html.it/news/2514.htm.

realiza manipulación alguna. Pero, poco antes del juicio por la demanda de American Blind, Google había decidido eliminar algunas AdWords adquiridas por los activistas de Oceana 26[26]. La «culpa» de Oceana había sido la publicación de denuncias ecologistas de la actividad de un gran inversor de Google, la Royal Caribbean Cruise Lines, utilizando como *keywords* en sus anuncios palabras como «cruise vacation» o «cruise ship», utilizadas normalmente para buscar páginas reservadas a las vacaciones en barco o similares. Se aduce como motivo oficial que Google es un canal neutro y que no quiere ser vehículo de campañas de propaganda desacreditadora de otras empresas. Evidentemente, en este caso la libertad de expresión no era tan importante.

Además, el mismo día de la audiencia sobre el caso de American Blind, en el distrito de San José, donde se celebraba el juicio, los resultados de Google aparecían misteriosamente diferentes de cualquier otro lugar del mundo. Por primera vez, Google manipulaba los resultados con fines distintos de aquellos del «mejor resultado de una búsqueda». El veredicto positivo para Google del tribunal respecto al caso Geico (muy similar en todos los sentidos al de American Blind) no suprime esta realidad.

El caso más reciente y conocido es el de la entrada en China del coloso de la búsqueda. Para conquistar el creciente mercado chino, Google se ha sometido por primera vez a las peticiones de censura, haciendo inaccesibles a los usuarios que efectúan búsquedas desde el territorio chino sitios señalados por las autoridades. Aunque ya en 2002 un estudio de Harvard había revelado que Google censuraba 113 sitios, en las versiones locales francesa y alemana (Google.fr y Google.de). Google contestó que era verdad, pero que habían sido borrados bajo indicación de entes gubernamentales, autoridades policiales y otros sujetos, y tan solo después de un atento análisis de sus contenidos: muchos sitios eran de índole racista, otros de religiones fundamentalistas. Alguien polemizó sobre el hecho de que la tan publicitada transparencia de Google empezaba a desmoronarse: había que notificar a los usuarios la aplicación de una censura que, de otra manera, podía quedar «oculta»; otros anotaron que la responsabilidad no era de Google sino de los ordenamientos jurídicos que contemplan la posibilidad de llevar a juicio una empresa por la presencia de tan solo un link en sus paginas[27]. En aquel caso, es evidente que Google, para evitar posibles consecuencias legales, optó por evitar las posibles consecuencia legales y decidió eliminar los links evaluando caso por caso. Hay que apuntar que el final del derecho de link puede acabar pesando sobre las libertades digitales en una medida cada vez mayor. ¿Quién decidirá qué es licito censurar? ¿Una súperautoridad más, a lo mejor de carácter supraestatal?, ¿el derecho del más fuerte, o sea, en el sistema del mercado, el derecho del que paga (o vale) más?, ¿o los chantajes de los fundamentalismos locales, sobre todo de matriz religiosa, que amenazan con venganzas cada vez que se condena su particular visión del mundo por algún «sitio subversivo»? El alcance del problema es tan amplio como la importancia de la libertad de expresión y seguramente no se puede solucionar en los tribunales. El ruido levantado por el caso chino, sin embargo, está relacionado con una censura ordenada por un Gobierno; a pesar de esto, Page y Brin no han cambiado de idea, conscientes de la importancia de un mercado dirigido a la cuarta parte de la población mundial.

Según Google, pronto el mundo será un enorme índice en el cual los recursos digitales se corresponderán perfectamente con la definición de la realidad circundante; cada índice podrá ser recogido y ofrecido de la mejor forma a través de

26 El grupo sin ánimo de lucro Oceana lucha por la calidad de las aguas: www.oceana.org.

27 La versión integral del informe firmado por Jonathan Zittrain y Benjamín Edelman se encuentra en la dirección http://cyber.law.harvard.edu/filtering/google/.

un algoritmo en los resultados de una búsqueda: y Google está en la posición perfecta para ser la herramienta de gestión de este índice.

Dejando a un lado la obvia y banal objeción de que los mundos reales y los digitales, en tanto que entrelazados, no se corresponden nada en absoluto, desde un punto de vista técnico no existe el algoritmo perfecto cuando se habla de encontrar las informaciones en la Red; sobre todo, nada es neutral en el ámbito tecnológico, y menos aun cuando se gestionan los datos de individuos reales en Internet.

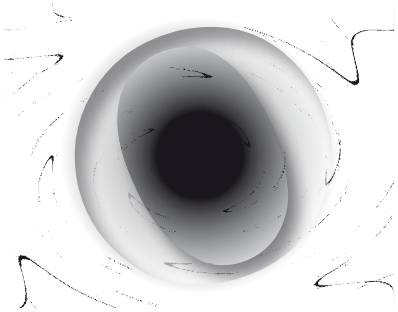
Independientemente de las alianzas que se lleven a cabo, la convergencia tecnológica que se está realizando nos parece una dirección nueva de la visión de Google como único punto de acceso, gestión y mediación de los datos digitales. La distopía de Google como aspirante a Gran Hermano se perfila peligrosa y fascinante, como todo choque histórico: la Web es el nuevo territorio de competición salvaje en pos de la afirmación de un nuevo estándar de comunicación. Un estándar paradójicamente «personalizado» con servicios y propuestas diversificadas según los gustos de los usuarios: ya desde hace algunos años la palabra clave es «personalización en masa» (*mass customization*). Es cierto, un oximorón que, sin embargo, revela la importancia de la apuesta en juego, un verdadero desplazamiento de paradigma, desde el consumismo en serie de la industria de masa al consumismo personalizado, que se vende como «libertad de elección». Más allá de las fáciles retóricas, la vía de fuga solo puede ser un nuevo rumbo en nuestras elecciones: no se trata de utilizar o no Google y sus servicios, sino de elegir una manera distinta de colocar informaciones en Internet y aprender a relacionarlas de una forma nueva, con recorridos nuevos y más interesantes para todos[28].

28 En este sentido, una fuente rica en referencias y propuestas, indispensable también para la redacción de este texto, es Ferry Bite y Claudio Parrini: *I motori di ricerca nel caos della rete-kit di sopravvivenza*, shake, Milán, 2001.

Desde hace unos años Google está comprendiendo a su costa (y, por supuesto, sobre la piel de sus usuarios) que la pureza no es una cualidad del mundo y tampoco del mundo de los negocios, que la bondad total es generalmente una tontería, más aún para una sociedad que busca ganancias y, sobre todo, que la neutralidad es un vía difícil de recorrer en la guerra de los motores de búsqueda. Además sería útil recordar que las naciones neutrales desde siempre, como Suiza, también son aquellas armadas hasta los dientes. Vamos a ver cuáles son las armas «buenas» de Google, las que le han permitido ser un fenómeno de relevancia mundial.

2

BeeGoogle



La huida de cerebros a Google: la guerra por el control de la Web

«¡Mataré a ese cabrón de Google! Eric Schmidt es un cabrón y lo enterraré, como ya hice»[1], así es como explotó en mayo de 2005 Steve Ballmer, consejero delegado de Microsoft, al conocer la noticia del fichaje por parte de Google de Kai-Fu-Lee, un alto directivo de la casa de Redmond y su hombre clave en China (Kai-Fu-Lee había desarrollado el motor MSN Search para cien millones de usuarios chinos). Una invectiva dirigida con toda evidencia contra el actual consejero delegado de Google, ex top manager de Sun Microsystems y de Novell, empresas a las que Microsoft ya se había enfrentado en el pasado tanto en el mercado como en los tribunales. Kai-Fu-Lee había sido nombrado responsable de un laboratorio de investigación abierto en los alrededores de Shanghai.

---

1 Declaración bajo juramento de Marc Lucovsky, empleado de Microsoft, durante una vista ante el tribunal del Estado de Washington, en el caso Kai-Fu-Lee; Lucovsky destacó también una presunta «sed de venganza» por parte de Microsoft. Las noticias proceden de muchos artículos de *Punto Informatico*, http://punto-informatico.it/.

Microsoft interpuso de inmediato una demanda contra el ex empleado y sobre todo contra Google, acusando a Kai-Fu-Lee de haber violado, gracias a la determinante ayuda del rey de los motores de búsqueda, unos reservadísimos acuerdos contractuales estipulados con el coloso de Redmond. Habiendo sido director ejecutivo, Kai-Fu-Lee conocía secretos comerciales e industriales y, según los abogados de Microsoft, no tendría reparo en utilizar tecnologías y *know-how* socioeconómico para maximizar las ganancias de la competencia. El hombre del escándalo costaría caro, aproximadamente dos millones y medio de dólares al año de «sueldo» y 20.000 acciones de la empresa. Cifras exorbitantes que dan la medida de la altísima apuesta en juego, no solo en el mercado chino.

La larga contienda judicial entre los dos colosos concluiría con un pacto bilateral en diciembre de 2005, justo un mes antes de acabar ante el juez: un acuerdo privado absolutamente confidencial. Puede que hayan sido pagadas grandes sumas de dinero, o quizás Microsoft haya obligado a Kai-Fu-Lee a no revelar nunca ningún tipo de secreto.

Esta historia solo es una de las más curiosas y reveladoras de una tendencia en marcha desde hace unos años: de hecho, Kai-Fu-Lee solo es el último directivo que se ha pasado a Google, «la sociedad que más se parece a Microsoft», como ha tronado Bill Gates, cada día más dudoso entre demonizar a los dos estudiantes-prodigio —reforzando indirectamente su notoriedad de protagonistas «buenos y generosos» del mundo de las tecnologías de la información— o fingir que no existen, que no son interlocutores de suficiente nivel.

En realidad, Bill Gates sabe de sobra que las migraciones de mánagers, particularmente en el sector tecnológico, conllevan la asunción del *core business* («corazón del negocio» o actividad principal) de la sociedad desde la cual el manager procede: Microsoft ha usado la misma estrategia con sus rivales en diversas ocasiones. La táctica comercial de robar los secretos industriales, mecanismos de producción y gestión de los recursos, mediante el fichaje de figuras punteras de las

sociedades adversarias ha sido siempre un recurso de la competencia industrial; y cada vez es más frecuente en la era de la economía de la información.

La elección del management por parte de Brin y Page es un claro indicador de los objetivos de Google: llegar a ser la plataforma de la Web más completa y personalizable, promoviendo una adaptación de los servicios a las necesidades de cada usuario sobre la base de un ingente cauce de informaciones. En esencia, Google aprieta el acelerador para catalogar todo tipo de información digital, desde los sitios web a los foros de discusión, de las imágenes a los vídeos, de los e-mails a los blogs, sin que se entrevea límite alguno en el horizonte; esto significa entrar en guerra con Microsoft, que de momento sigue siendo a pesar de todo el mayor competidor en el sector, con su navegador Internet Explorer, su portal MSN, su servicio de correo electrónico Hotmail, etc.

Los intereses de las dos sociedades se están solapando de manera creciente: las dos tienden a ser mediadoras indispensables para toda actividad informática. Microsoft ha conseguido imponer su dominio gracias a los sistemas operativos y los programas para oficina, que actualmente son un estándar en el mundo del trabajo y en el *home-computing*; por su parte, Google se impone hace tiempo como mediador global de los servicios web, tanto en la búsqueda de informaciones (que es su *core business*) en todas sus variantes, como también en servicios adjuntos como el correo electrónico. Esquematizando, podríamos decir que Microsoft está en una posición de dominio desde hace años gracias a productos que llevan a servicios, mientras que Google intenta dominar a través de servicios que necesitan de productos de soporte.

La competencia por tanto se fundamenta en las elecciones de los usuarios y en los estándares futuros que Google intenta imponer. Utilizar programas vía Web limitando solo al navegador los requisititos para acceder a los servicios significa sustraer mercado a quien siempre invirtió en el producto y en la creación de nuevas arquitecturas informáticas basa-

das en la Web. Lo mismo vale para los mercados económicos: se pasa de una economía al por mayor (Microsoft), donde se intenta vender licencias del mismo programa/producto, a una economía totalmente personalizada para cada usuario, donde los productos se pueden encontrar en la Red.

Largas colas en las redes: las entrañas de la economía de la búsqueda, Google versus Microsoft

Google aprovecha, según la fundamental argumentación del extenso texto de John Battelle[2], el ascenso de la «economía de la búsqueda». En su ensayo *The Search*, Battelle, periodista y uno de los fundadores de WiRed, sostiene que el futuro del comercio online estará ligado a las búsquedas personalizadas pagadas por los usuarios. Google, que gestiona la mayor base de datos de las «intenciones de búsqueda» de los cibernautas, se encuentra en la posición ideal para satisfacerlas, gracias a su Red capilar, constituida por una plataforma publicitaria excepcionalmente eficaz (AdWords) y por un network de anunciantes que cuenta con millones de sitios en todo el mundo (AdSense). La apuesta de Google es la de satisfacer toda necesidad/deseo que los usuarios expresen en sus búsquedas, ofreciendo servicios para el nuevo consumismo «a la medida del individuo»: cada usuario/cliente encontrará justo lo que busca, el producto estudiado aposta para él. El ejemplo más conocido de «personalización en masa» es el de los servicios de comercio online ofrecidos por Amazon, capaz de ganar mucho más de la venta de copias sueltas de un libro o de un CD que de los picos de ventas de cientos o miles de copias de best-sellers. Los muchísimos clientes que compran online libros no tan populares representan una gran cantidad de «eventos» que suceden raramente o incluso una sola vez;

2 John Battelle, *The Search How Google and Its Rivals Rewrote the Rules of Business and Transformed Our Culture*, Portfolio Handcover, 2005.

llegar a satisfacer estas «búsquedas personalizadas» representa el punto fuerte de la distribución de Amazon. De hecho, a una librería tradicional, basada en almacenes, stocks, pedidos mínimos, le es imposible ofrecer la disponibilidad de millones de títulos como hace Amazon; las mayores ganancias proceden por tanto de las novedades y de los títulos más vendidos. El suministro de una sola copia para satisfacer la petición de un solo cliente resulta antieconómica para una librería tradicional, mientras que para Amazon es posible y rentable, aprovechando la «economía de la búsqueda» de los mercados online.

Este tipo de mercados se denomina en jerga económica de «larga cola» (*long tail*)[3].

La teoría estadística de las largas colas se remonta al menos a la «distribución de Pareto»[4], según la cual «pocos eventos se producen con extrema frecuencia, y muchos eventos se producen raramente». Una distribución parecida de los eventos se describe, a nivel estadístico, con una hipérbole en la que la «larga cola» está representada por la multitud de «eventos», singularmente casi irrelevantes, pero en su conjun-

3 Las largas colas: las implicaciones comerciales de la larga cola (*long tail*) —en el contexto del comercio electrónico este concepto se ha vuelto de uso común— fueron analizadas por Chris Anderson, redactor jefe de *WiRed*, en un artículo de octubre de 2004 escrito para su revista. Anderson publicó en 2005 el libro *The Long Tail*. Como Battelle, Anderson ha mantenido un blog sobre la génesis de su libro en www.thelongtail.com y el término se ha vuelto de uso corriente. Las largas colas no son un invento de Anderson: se trata de un concepto familiar para cualquiera que haya hecho un examen de estadística. Hay distribuciones estadísticas muy comunes, cuyos gráficos describen un modesto número de eventos que ocurren con frecuencia, y un gran número de eventos (la larga cola) que suceden raramente. En su conjunto, los eventos raros pueden ser más numerosos que los normales: http://www.technologyreview.it/index.php?p=article&a=551. Un artículo sobre la larga cola en Italia y el *webmarketing*: http://www.minimarketing.it/2008/07/teorie-che-non-lo-erano-la-coda-lunga.html.

4 Vilfredo Pareto (1848-1923) fue sociólogo y economista. Una buena introducción se encuentra en Wikipedia, http://es.wikipedia.org/wiki/Vilfredo_Pareto.

to numéricamente muy elevados. Matemáticamente, una distribución de larga cola sigue la denominada «ley de potencia»[5].

La estrategia ganadora en los mercados de «larga cola» por tanto no es la bajada de los precios de los productos más populares, sino la ampliación de la oferta, que permite mayores ganancias a través de la venta de productos distintos en pocas cantidades.

En términos comerciales, la mayor rentabilidad procede de las pequeñas ventas. En efecto, la mayor parte del negocio en la Red se basa en la larga cola: el propio Google produce beneficios vendiendo publicidad a bajo coste a millones de personas en forma de banners textuales (*text ads*), mucho más que vendiendo pocos espacios de elevado coste a importantes anunciantes.

Battelle se interesa por la aplicación de la investigación a mercados no explotados: en el caso de Google, la enorme cantidad de datos a disposición para realizar búsquedas ha permitido, según él, la explotación comercial de las largas colas. En el contexto del *e-commerce* las largas colas tienen tres implicaciones. En primer lugar, a través de Internet los productos con poca demanda pueden, colectivamente, crear un mercado superior al del pequeño grupo de los artículos superventas. En segundo lugar, Internet da vida a una proliferación de vendedores, así como favorece una proliferación de mercados (es el caso del sistema de subastas online de e-Bay). Por último, gracias a la investigación, se hace posible el paso de los mercados de masa a los de nicho.

Esta última tendencia se basa sobre la formación de grupos espontáneos, un fenómeno muy difundido en las redes. En Internet, hasta los grupos numéricamente más relevantes no son masas homogéneas de personas, sino variadas comunidades de usuarios unidos por una pasión, un interés o un fin común. La oposición entre masas y nichos no tiene por tanto mucha relevancia para la detección de segmentos de mercado

5 Véase: A. L. Varabais, *Link*, Einaudi, Torino, 2001.

útiles. Al nivel comercial esto provoca el nacimiento de sitios que proponen *e-commerce* de productos ligados a tipologías muy especificas de usuarios, que fuera de la distribución online no constituyen un mercado nada interesante: por ejemplo, camisetas con logos para cybermaniáticos, relojes que marcan la hora en código binario, ordenadores luminosos y otros objetos dedicados a tecnófilos. La amplitud de la oferta compensará la escasez de la demanda, repartida entre un número elevadísimo de productos altamente personalizados.

En tal escenario es probable, como pone en evidencia el atinado artículo de Charles H. Ferguson[6], que Google y Microsoft luchen por el control de la catalogación, de la búsqueda y extracción de todas las informaciones, en todo tipo de dispositivo digital. Microsoft está invirtiendo mucho en los servicios web. En noviembre de 2004 lanzó una versión beta (experimental) de un motor de búsqueda proyectado para contestar a las peticiones formuladas en lenguaje corriente y para ofrecer resultados personalizados sobre la base de la localización geográfica del usuario[7]. En febrero de 2005 el motor MSN search fue mejorado y con él es posible también buscar dentro de Encarta, la enciclopedia multimedia de Microsoft; por el momento, la consulta está limitada todavía a dos horas y, en el curso de la navegación, un contador de minutos va actualizando el tiempo que queda. En suma, el gigante Microsoft ha decidido desarrollar un sistema propio para las búsquedas en la Web y en los ordenadores personales sin pasar por Google, que desde hace ya varios años es el triunfador de la búsqueda (tan solo Yahoo! puede presumir de una influencia comparable).

En su conjunto, los mercados ligados a la economía de la búsqueda parecen mucho más amplios que el actual mercado de los servicios de búsqueda. Microsoft anda con retraso, es cierto, pero podría poner en marcha estrategias feroces e in-

6 Véase el artículo en italiano de Charles H. Ferguson, publicado en *Technology Review*: http://www.technologyreview.it/index.php?p=article&a=340.

7 El motor de búsqueda de Microsoft: http://search.msn.com.

alcanzables por Google: por ejemplo, podría invertir en pérdida y ofrecer acceso gratuito a su motor de búsqueda, después de haberlo integrado en el navegador Internet Explorer; o, quizás, vender publicidad por debajo del coste a los anunciantes para arrebatar liquidez a su competencia. Mientras tanto, el sistema operativo desarrollado en Redmond, Windows Vista, estaba llamado a ofrecer herramientas de búsqueda innovadoras, expectativas que no se han visto cumplidas. Por otro lado, Microsoft llevaba un fuerte retraso respecto a Netscape (el primer navegador descargable gratuitamente de la Red) en la batalla por el dominio sobre los navegadores, pero sin embargo Explorer ha conseguido desbancarlo, y no ciertamente por ser el mejor navegador del mercado. Si Microsoft tiene de su parte una larga experiencia, una posición de predominio en el mercado y un enorme poder financiero, también Google tiene unas buenas cartas para jugar. Es la encarnación de la joven empresa en expansión, se ha construido una fama de compromiso con la investigación y dedicación a la excelencia técnica, predica el culto de la velocidad en satisfacer las peticiones de los usuarios con interfaces sencillas y eficaces: se ha impuesto porque técnicamente es mejor que los demás motores de búsqueda. En la competición por el control de la Web, Google parecería tener una ligera ventaja; pero no hay que olvidar que la extensión de las actividades de Microsoft no tiene comparación: no se limita a la Web, sino que se extiende a todo el ámbito de las tecnologías de la información, desde las herramientas con Windows o MS Office a los contenidos con Encarta o la búsqueda de alto nivel con plataformas como dotNet, etc. Considerando la apuesta en juego —el dominio del acceso a cualquier información digital y las ganancias derivadas—, una convivencia pacífica parece cuando menos improbable. Por ahora, Google ha postergado el choque, pero éste no puede tardar en producirse.

La guerra de los estándares

Continuamos siguiendo a Ferguson en su razonamiento: el escenario que se abre es el de una guerra de los estándares. A día de hoy, hay tres actores involucrados: Google, Yahoo! y Microsoft. La industria de la investigación, como resulta también del análisis de Battelle, está en fuerte crecimiento. La tecnologías emergentes o en vía de consolidación, como por ejemplo el *streaming* de audio y vídeo a través de la banda ancha, el VoIP (telefonía vía Internet: piénsese en Skype y en el servicio homólogo de Google, Gtalk) y la mensajería instantánea generan enormes cantidades y tipologías de datos que aún esperan ser indexados, accesibles mediante toda una variedad de nuevos dispositivos electrónicos (de los PDA a los móviles, y de los lectores de documentos audiovisuales de nueva generación a los navegadores vía satélite, etc.), cruzados entre sí para satisfacer las peticiones de los usuarios y, en última instancia, explotados por una publicidad invasiva de todo tipo. Para poder cooperar recíprocamente y con muchos otros sistemas se necesitarán nuevos estándares, difíciles de introducir de forma indolora.

Las guerras entre los estándares se producen porque el mercado de la tecnología de la información necesita de idiomas comunes para poder gestionar el aumento de complejidad. El valor de la información reside en su difusión; pero difundir e intercambiar textos es más simple que difundir e intercambiar audio, o incluso vídeo: una mayor cantidad de datos requiere redes de comunicación más potentes y cambios en las modalidades de gestionar las informaciones. Históricamente, el control propietario sobre un importante estándar tecnológico ha producido siempre enormes riquezas[8].

8 El predominio arquitectural es equivalente a fabricar moneda y, gestionado de forma adecuada, puede durar para siempre. La arquitectura de los ordenadores *mainframe* IBM fue introducida en 1964; Intel desarrolló su primer microprocesor en 1971; el primer sistema operativo Microsoft fue presentado en 1981; Cisco Systems vendió su primer router en 1986. Nin-

A veces, como en el caso de la Web, la adopción de un estándar abierto y no propietario, como es el protocolo HTTP, lleva a una situación ventajosa para todos. Frecuentemente sin embargo, no se imponen las «mejores» tecnologías, sino que «es más importante contar con una estrategia eficaz. De todas formas los ganadores tienden a connotarse por algunas características importantes. Suelen vender plataformas universales, independientes del hardware, como los sistemas operativos Microsoft, en lugar de soluciones hardware y software fuertemente integradas, como los sistemas Apple y Sun. Las arquitecturas ganadoras son propietarias y difíciles de imitar, pero son también extremadamente "abiertas"; o sea, que ponen a disposición interfaces accesibles públicamente para el desarrollo por parte de los programadores independientes y los usuarios finales. De esta forma, la arquitectura puede alcanzar todos los mercados y crear un punto de enlace y de bloqueo: en otras palabras, empuja a los usuarios hacia aquella arquitectura específica, haciendo casi imposible el paso a un sistema de la competencia sin grandes dificultades ni gastos»[9]. Se impone un estándar cerrado para explotar una situación de monopolio.

Un ejemplo que aclara la lucha por la hegemonía sobre los estándares es el reto que enfrenta a Gtalk y Skype. En estos momentos Skype ha conseguido una posición de monopolio en la gestión del VoIP[10] para el mercado doméstico; pero es posible que no hayan sido evaluados atentamente los tiempos de asimilación de las tecnologías abiertas dentro de las comunidades de desarrollo. Hasta hace poco, Skype era la única solución realmente funcional para cualquiera que

guna de estas tecnologías está dando señales de agotamiento y cada una de ellas ha generado cientos de miles de millones de dólares de facturación.

9 Charles H. Ferguson, *ob. cit.*, *Technology Review*.

10 VoIP: *Voice over IP* (voz mediante Protocolo Internet). Tecnología que hace posible realizar una conversación telefónica mediante conexión de Internet u otra red que utilice el protocolo IP, en lugar de pasar a través de la línea normal de transmisión telefónica. [N. del E.]

quisiera telefonear vía Internet, aun no teniendo ningún conocimiento técnico; pero paulatinamente las cosas están cambiando. Las tecnologías propietarias de Skype podrían ser superadas por Gtalk, basado totalmente en tecnologías libres (particularmente en el protocolo de comunicación jabber) con librerías de desarrollo concedidas bajo licencias copyleft, que favorecen la confluencia de energías en el proyecto pues animan a los programadores a potenciar la red VoIP de Google. En este caso la elección del open source podría revelarse ganadora para quitarle el predominio a Skype, la que, por su parte, naturalmente podría decidir hacer públicos sus códigos para mantener sin cambios la situación. El hecho de elegir adoptar tecnologías y plataformas propietarias, que cierran el acceso (concediendo a lo mejor interfaces públicas de desarrollo), o abiertas es una parte fundamental de la estrategia para el control de la Web y de la economía de la investigación.

El acceso al mercado de la investigación ya ha sido bloqueado (en jerga económica *locked in*). De hecho, es imposible que una empresa *start-up* pueda competir ahora con Google o Yahoo! para la indexación de miles de millones de páginas web: la inversión para la maquinaria necesaria sería prohibitiva, aun cuando se dispusiera de un mejor algoritmo para su propio *spider*. De todas formas muchos aspectos colaterales, sobre todo la interfaz entre los distintos sistemas de búsqueda, dejan amplios espacios a posibles innovaciones «críticas» a un coste relativamente bajo. Es el caso de las librerías informáticas, pequeñas parcelas de software que permiten la conexión entre sistemas heterogéneos, puentes de traducción de un sistema a otro, las cuales —junto con las metodologías de integración entre distintos dispositivos y de intercambio de informaciones y resultados de búsqueda— podrían ser desarrolladas por pequeños investigadores independientes, en lugar de las grandes empresas involucradas[11].

11 Véase Ippolita, *Open non é Free*, Eleuthera, Milán, 2005, cap. V.

Más adelante profundizaremos en estos aspectos ligados a las interfaces y a las librerías.

Por ahora es importante subrayar que ninguno de los actores involucrados tiene una posición de predominio absoluto, y obviamente esto es un bien para todos: intentad imaginar qué significaría un monopolio total del mercado de la investigación conseguido por un sujeto privado cualquiera a través de la imposición de hecho de un estándar. Las primeras y obvias consideraciones habría que hacerlas a propósito de la privacidad: ¿de quién serían los datos indexados con los que se efectuasen las búsquedas y se generasen ganancias astronómicas? Además, puesto que hoy en día, simplemente tecleando el nombre de una persona en Google, se accede a una impresionante cantidad de datos, en un futuro no lejano, en el que la cantidad y la calidad de las informaciones —y sobre todo la posibilidad de llevar a cabo búsquedas adicionales sobre datos heterogéneos— mejorará drásticamente, el control sobre los individuos podría volverse infinitamente más agobiante y total, cruzando datos sensibles, fichas médicas, conversaciones telefónicas, informaciones del ADN. Google podría ser el punto de acceso privilegiado al panóptico[12] digital: veamos, por tanto, cuáles son sus mejores armas en esta verdadera guerra por el control de las redes.

Arma n.º 1: GooglePlex, el capitalismo blando

La hagiografía corriente de Google cuenta con deleite el impresionante crecimiento que ha llevado a Brin y Page desde el

12 El panóptico auspiciado por Bentham y analizado por Foucault encuentra en el proyecto Echelon una de sus realizaciones más escalofriantes. Sin dejar lugar a paranoias complotistas, destacamos entre los muchísimos volúmenes y artículos un texto breve introductorio de Alex Jones y Paul Joseph Watson, «Il panottico, una prigione di sorveglianza per l'umanitá», 20/01/2006, http://www.ariannaeditrice.it/articolo.php?id_articolo=1878; versión original en http://www.prisonplanet.com/articles/january2006/110106thepanopticon.htm.

cuchitril de Stanford al garaje de Menlo Park —que un amigo había alquilado a la recién nacida sociedad Google Inc.—, el despacho en la University Avenue de Palo Alto, hasta el traslado a Googleplex, los actuales despachos de Mountain View, California. Entre 1998 y 2000 los dos amigos siguen difundiendo la misma cultura universitaria, desarrollando una cultura de empresa hecha de innovación, creatividad y sacrificio. La dedicación universitaria para la investigación aplicada al negocio: ésta es la clave del éxito. Googleplex atrajo desde el comienzo a legiones de colaboradores: fue recreada la atmósfera del campus universitario americano, donde estudio, perseverancia, deporte y ocio se complementan y fusionan: si un ambiente lleno de comodidades y relajado favorece la creatividad entre los estudiantes, por supuesto que también podía favorecer la productividad de los trabajadores. El espíritu académico de fraternidad, de la élite universitaria que trabaja duramente para conseguir siempre lo mejor, es lo que se desprende de las historias que circulan acerca de Googleplex: se dice que se llegaban a cerrar secciones enteras del parking dos veces por semana para dejar espacio para los partidos de hockey sobre ruedas; hoy más que nunca en los despachos abundan todo tipo de artilugios y, sobre todo, destacan las lámparas de lava con sus emulsiones coloreadas; la informalidad elegida como sistema, con Larry y Sergey dirigiendo las reuniones semanales, a las que llaman TGIF (*Thank God It's Friday*: gracias a dios es viernes), y en las que decenas de empleados se juntan en el *open space* situado entre las mesas de trabajo.

Ciertamente tal atmósfera de trabajo alentó el espíritu de grupo y aceleró el intercambio de ideas. Googleplex parece ser casi más el lugar para expresar las propias pasiones de investigación que un puesto de trabajo; seguramente no es un puesto trabajo cualquiera, a pesar de las ya enormes dimensiones. Y eso que la organización estilo «campus», en EE.UU., existe desde hace al menos treinta años: Microsoft y Apple, tan solo por mencionar dos ejemplos, adoptaron

desde siempre un perfil de este tipo. La mitología de Silicon Valley está repleta de historias acerca de la creatividad por encima de todo, de la importancia de la cooperación entre los colaboradores; en suma, de la necesidad de tener trabajadores felices de trabajar y considerar los objetivos de la empresa al mismo nivel que sus propios objetivos vitales, en lugar de trabajadores frustrados, aplastados por jerarquías agobiantes, atrapados por horarios de trabajo fijos, reglas rígidas, ambientes opresivos.

Quizás la novedad de Googleplex está en haber promocionado conscientemente, desde el inicio, la imagen de sí como lugar «diferente», «nuevo», «para las mejores mentes». Entrar en Googleplex es imposible si no se conoce a nadie dentro. En el interior está vetado incluso hacer fotos, al menos en teoría. Como si hubiera que defenderse del mundo exterior, poblado por feroces tiburones financieros y malvados depredadores que solo quieren explotar el talento de los Google-boys.

Todo el mundo quiere trabajar en Googleplex. Una lista no oficial de las amenidades que esconde es casi obligatoria: gimnasios de empresa, piscinas, comida gratis en los cuatro restaurantes (uno de ellos vegetariano), bebidas y tentempiés gratis por doquier (¡basta ya de la esclavitud de las maquinitas de pago, Google lo paga todo!); campos de voleibol y baloncesto, y espacios al aire libre para hacer deporte, monopatines de motor para moverse entre los edificios del campus-empresa. Y esto no es nada comparado con la guardería, la escuela infantil y las escuelas de primaria de la empresa para los hijos de los empleados, totalmente gratuitas, o el estudio dental —un camión adaptado a estudio dental móvil—; y eso, en un país como Estados Unidos, donde la instrucción y la sanidad son un lujo al alcance de pocos, son posibilidades aún más increíbles.

Los puestos de trabajo son espectaculares, el sueño de todo fanático de las tecnologías digitales: monitor estándar de 21 pulgadas LCD, pantallas de plasma por doquier, juegos y juguetes (modelos del tamaño natural de los personajes de Star Wars, artilugios tecnológicos de todo tipo). Omnipresentes, las lámparas de lava con sus colores fluorescentes.

Googleplex es un lugar de ensueño, sumergido en el verde, donde los horarios de trabajo son flexibles, donde todo parece posible. En resumen, Googleplex transmite la filosofía de Google, el estilo de vida de Google: por supuesto, hay también una tienda donde comprar todos los gadgets corporativos que se puedan imaginar, alcanzables online en un sitio habilitado para ello; en su mayoría se trata de himnos a la inutilidad y al exceso como todo gadget que se respete, pero muy eficaces en su propósito de hacer del sentido de pertenencia a la empresa una razón de orgullo. La camiseta y el gorro con el logo de la compañía para consolidar el espíritu de equipo son verdaderos anacronismos: el presente del condicionamiento blando es Googleplex.

Además, no solo Google, sino también empresas como Apple y Yahoo! ofrecen desde hace tiempo un verdadero catálogo de objetos de consumo, desde líneas completas de vestuario a todo tipo de accesorios *high-tech*, lectores MP3 y lápices USB, hasta objetos de decoración, todos oportunamente marcados con logos y lemas corporativos. El muy versátil negocio de Google ofrece pijamas para los recién nacidos, y también Google Mini[13]. Googleplex es la encarnación del capitalismo de la abundancia en la era del informacionalismo[14]: todas las informaciones del mundo ofrecidas a todo el mundo de forma gratuita. Ya no es tiempo de escasez: la abundancia y la disponibilidad de bienes (en

13 Google Mini es un programa que permite que los ordenadores internos y en red de una empresa o de un particular, y los servicios externos que se ofrecen a través de su web, utilicen parte de la tecnología Google. [N. del E.]

14 Sobre el concepto de «informacionalismo» véase en particular Manuel Castells, *La era de la información: economía, sociedad y cultura*, Alianza Editorial, Madrid, 1997 (última reedición, 2003).

este caso de informaciones) es total. No olvidemos, sin embargo, que al fin y al cabo la fuente de esta abundancia es casi solo la publicidad hecha principalmente con links textuales. Casi todo lo demás es *free,* libre, en un sistema que funciona perfectamente[15].

Dos de los más influyentes personajes de este capitalismo de la abundancia son Mike Moritz[16], galés, quien, además de Google ha financiado también Yahoo; y John Doerr[17], quien también tuvo importantes participaciones en Sun y en Netscape.

15 Hay muchas señales de gazapos de Google; parece obvio que se multipliquen los problemas conforme crecen el tamaño y los servicios. Además, la práctica de publicar versiones beta que los usuarios testan ampliamente acrecienta la posibilidad de disfunciones. Un ejemplo concreto: Marco Deseri en *Goggle Analytics*, http://www.deseri.eu/articoli/google-analytics-rallenta-il-sito.

16 Periodista del *Time Magazine* en los años setenta y ochenta, hoy se considera como un *veteran venture capitalist* y parece tener una receta para todos: ¿recrear hoy fenómenos como Google? Lo importante, dice, «es crear algo que sencillamente pueda ser útil para uno mismo». Artista, si no hubiera dado prioridad a los compromisos de trabajo, es hombre de marketing de la vieja escuela, pero siempre incisivo: las ideas correctas son aquellas que satisfacen a uno mismo y, por tanto, a los consumidores o, mejor aún, a los «usuarios».

17 John Doerr continúa también después de Google sus inversiones: últimamente ha apostado 16 millones de dólares en un proyecto llamado «Zazzle», cuyo *core business* es el mercado online adonde se pueden comprar y vender productos altamente personalizables. En la práctica, Doerr no quiere quedarse fuera de la «personalización en masa», en fase de actuación avanzada en Estados Unidos. Inspiradores del proyecto son personajes como Eric von Hippel, quien, «desde su cátedra de la Sloan School of Management teoriza en el MIT desde hace tiempo sobre la idea de que el próximo paso será exactamente ese: transformar la producción en masa fordista y postfordista en algo completamente distinto, a medida, creado y gestionado a través de la Web», http://www.bloggers.it/Marcioweb/index.cfm?idForum=News%20dal%20Mondo&blogaction=archive&file=blog_7_2005.xml.

Arma n.º 2: perfeccionar la estrategia de la acumulación

Después de una visita a Googleplex se entiende el motivo de la huida de cerebros a Google de la que hablábamos al comienzo del capítulo. Para un obrero o, mejor dicho, peón de las tecnologías de la información, trabajar en Googleplex es un sueño hecho realidad. En este sector cada vez son más los obreros explotados; un perfil profesional ejemplar sería el del *coder* independiente que desarrolla proyectos personales, a lo mejor publicándolos en SourceForge.net o Slashdot.org, y ofrece sus conocimientos en el mercado sin tener ningún tipo de protección sindical, contrato de categoría o similares, ya restos prehistóricos en nuestra época de la flexibilidad total. En Googleplex hasta podrá dedicar el famoso 20% del tiempo de trabajo a sus proyectos, pagado y motivado a hacerlo cada vez mejor.

Entre partidos de hockey y voleibol, perros que merodean por el campus/empresa y en los pasillos, reuniones alrededor de una mesa de ping-pong, ciertamente es difícil aburrirse. Ya que no es fácil fichar nuevos colaboradores a fin de mejorar incluso este tipo de atmósfera, Google se ha distinguido por sus ocurrencias en materia de fichajes. La más curiosa quizás sea el test realizado en julio de 2004: en algunas estaciones del metro de la ciudadela de Cambridge, en el estado de Massachussets (EE.UU.), y en la autopista 101 en California, aparecieron grandes vallas blancas con el texto:

{first 10 digit prime in consecutive digits of e}.com

o sea

{primer número primo de 10 cifras consecutivas en e}.com

El logaritmo natural requerido es el número 7427466391; en la dirección http://www.7427466391.com/ indicado en la

valla nos encontramos con una dirección IP de Google que nos pide que completemos una secuencia de números; una vez encontrado el número 5966290435, se siguen las instrucciones, utilizándolo como password para entrar en una sección del sitio http://www.linux.org (paso nada obvio) y desde aquí se nos rebota otra vez a Google, al sitio http://www.google.com/labjobs/, en donde nos piden que enviemos nuestro currículo: si hemos logrado resolver los enigmas, podríamos ser un buen fichaje para Google.

Google no atrae tan solo a los mejores técnicos, hackers y lumbreras geniales: pronto también los mánagers más cotizados del mundo de las tecnologías de la información se darán cuenta de sus potencialidades y competirán por entrar en su plantilla.

La estrategia de la acumulación de datos para la búsqueda, y de máquinas en red para almacenar una copia de los datos, procede al unísono con la acumulación de cerebros. Máquinas significantes, máquinas electrónicas y máquinas biológicas se acumulan en Googleplex para crear un estilo de vida, una especie de religión de la excelencia, encarnada por un evangelizador.

El manager que mejor encarna el estilo de la empresa, el dichoso «evangelista» de Google, entre tantos jóvenes, es un gran viejo de la Red, Vinton G. Cerf, inventor junto con Robert Kahn del protocolo TCP/IP. Las circunstancias merecen un pequeño paréntesis: en febrero de 2005 Google confirmaba que el ICANN[18], el organismo que supervisa el sistema de los nombres y de los dominios Internet, había dado el visto

18 Internet Corporation for Assigned Names and Numbers (Corporación de Internet para la Asignación de Nombres y Números) es una organización sin ánimo de lucro «responsable de asignar espacio de direcciones numéricas de protocolo de Internet (IP), identificadores de protocolo y de las funciones de gestión [o administración] del sistema de nombres de dominio de primer nivel genéricos (gTLD) y de códigos de países (ccTLD), así como de la administración del sistema de servidores raíz», según sus propias palabras (http://www.icann.org/tr/spanish.html). [N. del E.]

bueno para que la empresa de Mountain View pudiera entrar en el negocio del registro de los dominios. En septiembre de 2005 Google anuncia que Vinton Cerf ha sido nombrado «vicepresidente y evangelizador jefe de Internet para Google. Su misión consiste en identificar nuevas tecnologías y aplicaciones estratégicas para la empresa en Internet y en otras plataformas»[19]. Hasta ese momento, entre otros cargos, Cerf ocupaba un puesto en el Consejo de Administración de la ICANN. A diferencia del fichaje de Eric Schmidt como consejero delegado, y de otros top mánagers robados a la competencia, ésta es sin duda alguna una operación exclusivamente de imagen (se hace difícil pensar que Cerf trabaje de verdad en Googleplex, por muy divertido que pueda ser...).

Arma n.º 3: la imagen lo es todo, pero un poco de «filosofía» no viene mal

La imagen de Google no es tan solo el sitio y la interfaz[20] que han conseguido tanto éxito por su sencillez y velocidad; no solo es Googleplex, el mundo dorado de los prodigios tecnológicos; no solo es «ser buenos» y hacer dinero, conjugando estrategias comerciales descaradas con los guiños a la cultura académica y a las comunidades free software y open source, con financiaciones e iniciativas hechas a su medida.

La imagen es también y sobre todo la «filosofía» de la empresa, la explicación clara y meridiana, en píldoras, del Google-pensamiento. No se trata precisamente de investigaciones llevadas a cabo por amor al conocimiento y a la claridad, por esto quizás el término filosofía es un poco exagerado; en cualquier caso, online se pueden encontrar los diez puntos, los diez mandamientos del buen Google. Se inicia con una

19 Véase http://www.google.it/intl/it/corporate/execs.html#vint. Recordemos que en EE.UU. prácticamente cualquiera que no sea empleado es vicepresidente de algo.

20 Véase el capítulo V.

introducción que anuncia el estilo: «Nunca conformarse con lo mejor»; de hecho, según afirma Larry Page, el objetivo último es «el motor de búsqueda perfecto [...] capaz de comprender exactamente la petición del usuario y devolver exactamente lo que éste desea». Por lo tanto, Google no quiere ser para muchos: quiere ser para todos, satisfacer los deseos de todos, llevar la felicidad al mundo. Para lograrlo, se dedica con tenacidad a la innovación y a la investigación, siguiendo las directrices de sus «diez verdades»:

1. *Atención al usuario: todo lo demás viene después*. El crecimiento de Google ha sido fruto del boca a boca, que ha atraído a usuarios encantados por sus prestaciones, y no de agresivas campañas de comunicación. La publicidad no debe ser invasiva, sino útil al usuario.

2. *Es preferible dedicarse verdaderamente bien a una sola cosa*. «Google se dedica a la búsqueda. Con uno de los grupos de investigación más grandes del mundo dedicado exclusivamente a la solución de los problemas de la búsqueda, somos conscientes de lo que hacemos bien y de cómo podemos hacerlo mejor». Hemos visto que Googleplex se configura exactamente como un centro de investigación universitario. Los nuevos servicios añadidos al simple motor de búsqueda se ven como oportunidades ofrecidas a los usuarios para orientarse y encontrar las informaciones que desean sin perderse en masas cada vez más grandes de datos.

3. *Rapidez es mejor que lentitud*. «Google cree en la gratificación inmediata. El usuario quiere respuestas y las quiere ya. ¿Quiénes somos nosotros para no estar de acuerdo? Google es prácticamente la única sociedad del mundo cuyo objetivo manifiesto es procurar que sus usuarios abandonen el sitio de Google lo antes posible». Dos son las principales intuiciones y realizaciones que le han permitido a Google obtener «velocidad»: la puesta a punto y el perfeccionamiento continuo del algoritmo PageRank(™), que indexa sin pausa las redes; el uso de plataformas modulares fuertemente escalables conectadas entre sí (*clusters*). De todas formas, sería

oportuno reflexionar más en profundidad sobre la velocidad como panacea: a veces, también en el entorno de las nuevas tecnologías, la lentitud es una virtud[21].

4. *La democracia en la Web funciona*. «El éxito de Google está basado en millones de individuos que publican sitios web para determinar qué otros sitios ofrecen contenidos válidos». Ya sabemos que Google utiliza PageRank(™) para evaluar los sitios enlazados a una página web y asignarles un valor, basado en parte en los sitios conectados a éstos. La imagen de esta democracia electrónica es idílica: el resultado del ranking de Google sería un «índice de aceptación popular» basado en un algoritmo/ley electoral que permitiría a los usuarios/ciudadanos de la Red expresar sus preferencias/votos a través de los links ofrecidos y aceptados por otras páginas, y encontrarlas puntualmente expresadas en el posicionamiento de los sitios web votados. La comparación link/voto es por lo menos forzada y simplista, pues continuamente se introducen «correcciones» para calcular el ranking, «pesando» de forma diferente los links/votos. Seguramente el link/voto de un sitio porno contará menos que el link/voto de un sitio universitario. Aunque sea al menos discutible que la cultura académica sea más popular que la cultura pornográfica... lo cierto es que, con el crecimiento de las informaciones, esta «democracia» está en expansión continua.

5. *No hace falta estar en la propia mesa de trabajo para necesitar informaciones*. «La sociedad actual es cada vez más móvil y está menos dispuesta a estar vinculada a un lugar fijo. A través de los PDA, los teléfonos inalámbricos o sus automóviles, los usuarios desean que las informaciones vayan hacia ellos». La flexibilidad de los tiempos y de los espacios es un objetivo importante: la convergencia de los dispositivos electrónicos (TV, radio, teléfono, Internet...) en dispositivos móviles en miniatura es una oportunidad total para el mayor

21 Para una panorámica completa sobre el tema, véase Thomas Hylland Eriksen, *Tempo tiranno – Velocitá e lentezza nell'era informatica*, Eleuthera, Milán, 2003.

proveedor mundial de soluciones de búsqueda. La penetración en este mercado del futuro es estratégica, lo hemos visto hablando de la «guerra de los estándares»: más aún para Google, que no produce dispositivos electrónicos para los cuales proponer/imponer su software (Microsoft, Apple), sino interfaces de búsqueda. Cada nuevo dispositivo es un territorio por conquistar.

6. *Se puede ganar sin hacerle daño a nadie*. «Google es una empresa. Los ingresos generados por la sociedad proceden de la oferta a las empresas de su tecnología de búsqueda y de la venta de publicidad colgada en Google y en otros sitios web». Pero los anuncios son textuales y poco invasivos, los links de pago se adecuan a las búsquedas realizadas (AdWords); y los usuarios pueden transformarse en anunciantes de forma sencilla y a modo de autoservicio; si se trata de un gestor de sitios web, puede incluso aprovechar la red de Google a través de AdSense, publicando anuncios adecuados al contenido de sus sitios. «No seas malo», «no hagas daño a nadie» significa también «no hagas a los demás la publicidad que no te han hecho a ti» y, por supuesto, garantiza que el PageRank(™) no es objeto de venta. La confianza de los usuarios en la precisión de los resultados es la fuerza principal de Google, y no se puede sacrificar para ganancias a corto plazo. Se aprovecha para generar ingresos indirectos, de segundo nivel (generados por la publicidad).

7. *Siempre hay más informaciones de la que se pueda imaginar.* «Después de haber indexado más paginas HTML en Internet que cualquier otro servicio de búsqueda, los técnicos de Google han dirigido su atención a las informaciones no inmediatamente accesibles». Sí, Google acumula bases de datos heterogéneas: imágenes, mensajes de newsgroups (Usenet), números de teléfono y direcciones, informes financieros. Para ser el principal mediador de informaciones del mundo, no hay que poner límites a la acumulación.

8. *La necesidad de informaciones traspasa todo límite*. «A pesar de que la sede de Google esté en California, tenemos

oficinas en todo el mundo y nuestra misión es facilitar el acceso a las informaciones a las personas de todo el mundo». Cultura norteamericana, estrictamente académica, para todos. Pensar en grande, siempre y como sea: indexar todas las informaciones, ofrecer acceso a todo el mundo. Las traducciones son parte fundamental del universalismo de Google: desde el coreano a la jerga hacker o el klingoniano de Star Trek, del hindi al xhosa surafricano o el *pig latin* (o sea el alfabeto «farfallino» norteamericano), desde el zulú a la interlengua o el bork bork bork (una especie de falso ruso molvano...), todos han de poder acceder. Más de cien idiomas de interfaz. Dominios locales de más de cien naciones. Loable, pero un poco centralizador... operación quizás políticamente correcta y formalmente respetuosa de las minorías, aunque en realidad se trata de un *super-layer*, un nivel superior, una única interfaz que allana, homogeniza las diferencias y difunde el estilo de Mountain View.

9. *Se puede ser serio sin americana y corbata*. «Los fundadores de Google muchas veces han declarado que la empresa es "seria" solo por lo que atañe a la investigación. Ellos han edificado una empresa basada en la idea de que el trabajo debe ser un desafío y que dicho desafío tiene que ser divertido». En este mandamiento se compendia la organización de Googleplex como un campus universitario para maximizar la productividad: se afirma explícitamente que «el énfasis se pone en los resultados logrados por el equipo y en el orgullo de los hallazgos individuales que contribuyen al éxito global de la sociedad», y que «este ambiente extraordinariamente comunicativo favorece una productividad y una camaradería alimentadas por la conciencia de que millones de personas confían en los resultados de Google. Dad los instrumentos adecuados a un equipo de personas que quieren ser diferentes y lo serán». Quizás éste sea la mejor manera de aprovechar los creativos, haciéndolos al mismo tiempo entusiasmados defensores de la «experiencia Google».

10. *No basta con ser excelentes*. «Dar siempre más de lo esperado. Google no acepta la idea de que el objetivo final sea el de ser los mejores, sino que tiene que ser el punto de partida. Con la innovación y la reiteración, Google se dedica a mejorar algo que ya funciona bien para conseguir resultados inesperados». Evidentemente, para satisfacer rápidamente, cada vez mejor, los deseos de todos los usuarios del mundo, hace falta retrasar un poco, siempre un poco más allá, la satisfacción de los propios deseos. Desear el deseo de ser los mejores. En este contexto llegar segundos es todavía peor que no existir.

Por lo que a nosotros se refiere, eslogan por eslogan, preferimos: «dar asco, en una sociedad que obliga a la excelencia, es un claro deber moral»[22].

Arma n.º 4: Google y el open source

El arma posiblemente más compleja de Google es su estrategia de colaboración-explotación del mundo open source. La iniciativa Google Code (marzo de 2005) es un homenaje a la comunidad open source, de la cual los fundadores de Google se declaran «amigos y deudores». El motor más usado de la Web explica que el fin de Google Code no es promocionar el desarrollo de aplicaciones que aprovechan sus propias API (Application Programming Interface), que ya disponen de un sitio ad hoc, sino el de poner a disposición de todos los desarrolladores open source herramientas de utilidad pública. Los cuatro proyectos inicialmente publicados en Google Code nacen alrededor de programas realizados por los mismos desarrolladores de la sociedad para agilizar la creación, la optimización o el *debugging* del código. Los proyectos linkados en Google Code también están hospedados en SourceForge.net y se distribuyen bajo licencia BSD 2.0

22 SantaMariaVídeo, la TV que no transmite nada: http://santamariavideo.blogspot.com/2010/03/uscire-dallanonimato.html.

(su código puede utilizarse indistintamente dentro de otras aplicaciones open source o propietarias). Además, Google promete desde ya poner a disposición de la comunidad open source muchos otros softwares, fruto en buena medida de ese 20% de tiempo del que disponen los desarrolladores empleados por la sociedad para dedicar a proyectos personales.

No es casual que Google siga siendo un proyecto dedicado a reclutar desarrolladores open source a través del programa «Summer of Code» (el verano del código), una selección de proyectos dotados con un premio de 4.500 dólares para su desarrollo. Luego le toca el turno a «Google Earth» y, finalmente, como toda potencia que haya llegado a representar un estilo de vida, Google presenta el sueño: «moon.google.com», ¡Google en la Luna!

«En honor del primer desembarco en la Luna, el 20 de julio de 1969, hemos añadido algunas imágenes de la NASA a la interfaz de Google Maps para permitir que todo el mundo pueda visitar a nuestro vecino estelar. Buena navegación lunar».

Los movimientos de Google, «pseudomonopolistas» por lo que se refiere a los métodos y perspectivas, han repercutido inmediatamente sobre la competencia: hoy Google se está transformando en un gigante que ocupa todos los espacios de mercado; cada vez nuevos servicios estrangulan las sociedades *hig-tech* más pequeñas, a las que les cuesta fichar ingenieros y técnicos, con el riesgo añadido de ver imitados sus nuevos productos.

El lanzamiento sin interrupción de nuevos servicios y la financiación desde dentro de potenciales *spin-offs* (productos derivados) de los empleados bloquea, de hecho, el mercado en términos de innovación tecnológica. ¿Quién financiaría hoy un proyecto web, con la posibilidad de que en pocos días sea Google el que lo lance?

Google ha sabido representar, a los ojos de los observadores, pero también de los simples usuarios de Internet, una especie de paladín del progreso: a partir del motor de bús-

queda, concebido para ser entendido rápida y sencillamente por los usuarios, se han multiplicado las propuestas de servicios e ideas.

Con la opción open source la economía relacional puesta en marcha por Google deviene una *Weltanschaung*, una cosmovisión, acogida incluso como desarrollo sostenible, como capitalismo bueno que difunde la abundancia, una solución económicamente ética a las búsquedas de los individuos.

3

Google open source

Teoría y prácticas: *open* no es *free*

Free software[1] (software libre) y *open source*[2] (código abierto) son palabras que se utilizan a menudo como sinónimos para indicar códigos o porciones de códigos informáticos; sin embargo, por mucho que a menudo describan objetos idénticos, reflejan perspectivas radicalmente distintas. *Free software* es un término que nace a principios de los años ochenta por iniciativa de Richard Stallman: se refiere a la libertad del usuario para utilizar y mejorar el software. Más en detalle, se puede resumir en cuatro libertades fundamentales:

1) *Libertad de ejecutar el programa para cualquier objetivo.*

2) *Libertad de modificar el programa según las propias necesidades* (para que esta libertad tenga algún efecto en la práctica es preciso tener acceso al código fuente del programa, ya que aportar modificaciones a un programa sin disponer del código fuente es extremadamente difícil).

---

1 La filosofía empresarial del free software: http://www.gnu.org/philosophy/free-sw.html.

2 La definición de open source: http://www.opensource.org/docs/definition.php.

3) *Libertad para distribuir copias del programa gratuitamente o tras una remuneración*.

4) *Libertad para distribuir versiones modificadas del programa, y permitir de esta forma que la comunidad pueda disfrutar de las mejoras aportadas*.

Por otra parte, la expresión *open source* nace a finales de los años noventa, fundamentalmente por iniciativa de Bruce Perens y Eric S. Raymond, que en 1998 fundan la *Open Source Initiative*[3] (OSI); la referencia es la *Open Source Definition*, que a su vez procede de las *Debian Free Software Giudelines*, es decir, una serie de diez directrices que definen qué criterios legales debe de reunir una licencia para que se la considere efectivamente «libre» o, usando el nuevo término, *open source*.

Es evidente pues que, por un lado, el free software pone el acento en la libertad: «el software libre es una cuestión de libertad, no de precio»[4]. Por otra parte, el open source se ocupa exclusivamente de definir, en una perspectiva totalmente interna a las lógicas del mercado, cuáles son las modalidades mejores para difundir un producto según criterios *open*, es decir, abiertos. El free software tiene un sentido que va mucho más allá del mercado, aun sin excluirlo a priori; el open source existe, como han explicitado sus promotores, para adaptar un modelo preexistente (el *free*, en el sentido de «libre») al mercado.

3 El sitio de la Open Source Initiative: http://www.opensource.org/.

4 Incluso en artículos recientes, de gran nivel, es preciso notar esta confusión entre *free software* y *open source*, como si el segundo fuese una mejora del primero en perfecta sintonía y continuidad. El movimiento fundado por Stallman contiene, sin duda alguna, muchas limitaciones concretas, no solamente en las relaciones con el mercado sino también en la asunción de posturas demasiado ideológicas; sin embargo, no nos parece correcto trazar líneas evolutivas simples cuando en realidad el panorama es muy fragmentado y complejo. Véase, por ejemplo, «Economia delle Reti open source: storia e dinamiche del movimento del software libero», http://www.pluto.it/files/journal/pj0601/economia_reti.html.

Los hackers de Stanford

Ya desde hace algunos años el software open source se considera comúnmente fiable, capaz de prestaciones elevadas y con costes sostenibles; en la práctica, hoy en día, se le considera a menudo mejor que los software propietarios, precisamente porque permite aumentar la confianza de un producto en virtud de una metodología distinta de desarrollo, abierta y sometida a un amplio mecanismo de revisión.

Gracias a la realización de software estables y funcionales como navegador, programas para oficina, editor y completos sistemas operativos (GNU/Linux), casi todo usuario se ha dado cuenta o ha oído hablar de la existencia de los programas llamados genéricamente copyleft u open source.

«Open source» era el término adecuado, y necesario, para sustituir la expresión «free software». De hecho, en inglés la palabra *free* tiene el doble significado de libre y gratuito: esta ambigüedad lingüística no despertaba demasiado entusiasmo desde el punto de vista económico. La sustitución por *open* fue una fórmula estratégicamente ganadora para mantener las características de cooperación libre sin renunciar a la posibilidad de un uso más estrictamente comercial de los softwares.

En realidad, en esa época se estaba produciendo un cambio radical en el mundo de las comunidades digitales espontáneas, o sea, aquellas comunidades con las que se relacionan todos los que ofrecen una definición positiva de hacking.

Estas agregaciones eran (y aún lo son) extremadamente complejas. Nos referimos a un intersticio cultural fluido en el que se forman y colaboran estudiantes, profesores, investigadores, profesionales libres, policías y delincuentes, programadores pagados por sociedades de desarrollo, apasionados y muchas más tipologías de hacker.

El movimiento del free software estaba empezando a tener una confrontación cerrada con la economía de mercado. La batalla de la Free Software Foundation se centraba en la

difusión de la licencia GPL (*General Public License*), creada por el fundador de la FSF[5] Richard Stallman; esta licencia vincula el artefacto en modo «viral» a las cuatro libertades más arriba nombradas. Sustancialmente cada modificación aportada al código bajo licencia GPL tiene que mantener la misma licencia de uso, imposibilitando de esta forma el cierre del código: este mecanismo se conoce como «permiso de autor» (*copyleft*, juego de palabras con copyright). Surgían y se difundían entonces las primeras distribuciones del sistema operativo GNU/Linux.

La mezcla entre el método de desarrollo libre y la *net-economy* iba a determinar en los años sucesivos al 2000 la explosión de los productos open source y el desencadenamiento del debate político acerca de la patentabilidad del software, el copyright y la gestión ético-política de todo lo que actualmente se define como obra del ingenio humano.

La empresa Google, aunque no sea una productora directa de software, no se ha quedado al margen de la sacudida provocada por el open source; como otras empresas dinámicas e innovadoras, Google ha cooptado sus metodologías y las ha puesto al servicio de su «misión».

La contigüidad entre Google y open source es espacial y temporal: en 1998 en Stanford, precisamente mientras Brin y Page pulían la primera versión de su motor de búsqueda, emergían algunos importantes proyectos free software; recordemos por ejemplo SND y Protégé que, en ámbitos muy distintos, el del audio y la web semántica, iban a conseguir un gran éxito en el escenario digital.

En Stanford, la cultura hacker, de la que procede en última instancia el open source, se respira como algo familiar: no es pues un azar que nuestro dúo, que se había formado en aquellos años, haya manifestado siempre una cierta predilección por el desarrollo sobre la plataforma GNU/Linux.

5 La Free Software Foundation en español: http://www.gnu.org/home.es.html.

Si existen diferencias sustanciales entre open source y free software, hay también elementos en común y continuidad de perspectivas. Por simplicidad y corrección, hablaremos pues de «metodologías y prácticas abiertas» para indicar el común denominador que atraviesa el free software, el open source y la competencia de mercado en el mundo de las tecnologías de la información.

La primera característica de una comunidad *open* (y en este sentido también *free*) es la de poner en práctica un método de trabajo abierto a la colaboración de todo el mundo, o sea, capaz de aceptar sugerencias e interacciones espontáneas procedentes de toda tipología de sujeto implicado en la construcción del artefacto informático: programador, traductor o incluso simple usuario. Este procedimiento ha sido definido en la jerga hacker como método de «bazar» y su aplicación a gran escala se debe al desarrollo del Kernel Linux en los primeros años noventa, un proyecto nacido por iniciativa de Linus Torvalds y base de todas las distribuciones GNU/Linux[6].

La nueva técnica cooperativa propuesta por el *underground* digital ha vuelto del revés la ley de Brooks[7] que reglamentaba las comunidades de desarrollo de los proyectos informáticos hasta aquel entonces. Según la ley de Brooks, en la medida en la que crecen las complejidades aumentan de forma exponencial los errores y, por lo tanto, un proyecto al que contribuyen millares de desarrolladores tendería a ser un revoltijo de código inestable y lleno de bugs. En cambio, mediante la revelación de los códigos fuente, la circulación libre en Internet de la documentación, la cooperación y el *feedback* espontáneo de un número cada vez más elevado de sujetos en juego, las comunidades libres han demostrado que es posible una enorme mejora en la construcción de artefactos

6 Respecto a esto véase: Eric S. Raymond, http://biblioweb.sindominio.net/telematica/catedral.html.

7 La ley de Brooks según Eric S. Raymond: http://biblioweb.sindominio.net/telematica/catedral.html.

digitales, tanto desde el punto de vista del resultado como del proceso. Los softwares realizados de este modo se publican bajo una licencia GPL, que como hemos visto difunde de manera viral materiales copyleft.

La licencia GPL no prevé por lo tanto restricciones desde el punto de vista comercial; sin embargo, como el término free software resultaba demasiado radical al poner las libertades en primer lugar, así también la GPL ha sido sustituida por licencias edulcoradas respecto al contenido ético y político que el movimiento originario quería expresar. Este es el caso de licencias como la BSD (*Berkeley Software Distribution*), que no comporta ninguna restricción respecto al cierre de los códigos y por ello inhibe el efecto viral, porque partes del código no libre pueden ser integradas en el código libre. Entonces, una creación libre puede convertirse en propietaria. O la MPL (*Mozilla Public License*) y otras licencias desarrolladas exclusivamente para los nuevos productos open source.

La economía de mercado se convierte así en desarrollo sostenible y la *community* de los desarrolladores en el núcleo de una verdadera y auténtica *open society*[8], la quimérica sociedad abierta. Este imaginario está determinado no solamente por la adhesión moral que suscita la práctica de un desarrollo comunitario, sino sobre todo por la calidad superior de las aplicaciones, en aparente contradicción con la gratuidad de las competencias puestas en juego.

La era de la *open source economy*: competencia y buen rollo

El ingreso del open source en el mercado es, según algunos observadores, una de las consecuencias de la llamada «convergencia tecnológica», uno de los eslóganes que se ha

8 Para una profundización en este tema, remitimos a Ippolita: www.ippolita.net/onf/.

convertido ya casi en un paradigma de la era informacional: el acercamiento y la sinergia de varias tecnologías anteriormente consideradas extrañas, estudiadas y desarrolladas en ámbitos separados.

Ante estas transformaciones, a menudo demasiado rápidas, la creación de estándares abiertos ha creado un umbral en la «guerra de todos contra todos» del llamado «mercado libre»: *cooperate on standards, compete on solutions* (cooperar en estándares, competir en soluciones), es el lema de IBM, una de las principales empresas implicadas. Si incluso Big Blue decide cooperar, significa que el juego vale la pena...

Para muchas empresas el open source representa de hecho una de las pocas posibilidades para contrarrestar monopolios y oligopolios ya consolidados, para rehuir las dinámicas de competición clásicas sin inversiones enormes, para limitar los costes de desarrollo y reducir, pues, el «precio» de sus propios servicios.

Las empresas conocen desde hace tiempo y aprecian el valor de una dinámica reticular de desarrollo y alianzas: se sabe que el valor de una red es proporcional al cuadrado de las personas/nudos que conecta[9]. Una red cada vez más amplia comporta beneficios exponencialmente mayores.

9 Se trata de una ley matemática formulada a finales de los años setenta por Robert Metcalfe, estudiante de la Harvard University y luego fundador de la sociedad 3Com, además de ser pionero en el *networking* (e inventor del protocolo ethernet, aún hoy fundamental para las redes intranet). He aquí la traducción ad hoc del fragmento: «El valor de un network crece exponencialmente al número de los ordenadores conectados al network mismo. Por lo tanto, cada ordenador que se añade al network por un lado utiliza los recursos (o las informaciones o las competencias) del network y por otro aporta nuevos recursos (o nuevas informaciones o competencias) al network, aumentando así el valor intrínseco». De este principio general sigue que: 1) el número de relaciones posibles, o metarrelaciones o conexiones dentro de un network crece exponencialmente al número de ordenadores que se le conectan (de ahí el valor estratégico de los links dentro de una red); 2) el valor de una comunidad crece exponencialmente al crecimiento de los usuarios registrados en esa comunidad (de ahí el valor estratégico de las comunidades virtuales). Véase http://es.wikipedia.org/wiki/Ley_de_Metcalfe.

El open source parece ofrecer algunas garantías relevantes en el desarrollo de redes con alto valor añadido: por un lado, permite que el software siga siendo de alguna forma un bien «público» (adopta el desarrollo abierto y se vale de la comunidad de apoyo); por otro, mantiene muy baratos los costes de paso de un sistema a otro (los llamados *switching costs*), en especial, el paso de los modelos propietarios a los abiertos, sobre todo en el caso de los sistemas *legacy*, obsoletos. Los mayores costes en la adopción de nuevas tecnologías se deben a la formación de los usuarios, y no a la inversión en la compra de la tecnología misma; con mayor razón si se trata de excelentes softwares libres por un coste irrisorio. Pero el resultado más importante, cuyo precio es difícilmente calculable, es la creación de una imagen completamente nueva: para la propia empresa y los propios productos.

El éxito de las lógicas del software libre ha producido distintos intentos de aplicación de las mismas prácticas en otros sectores. Inevitablemente se han acuñado nuevas expresiones y se han formulados predicciones a menudo exageradas: *open laws*, *open science* e incluso la *open society* parecen al alcance de la mano. De hecho hoy la idea de sociedad open source se ha convertido casi en un paradigma de la nueva era dirigida a la búsqueda de un instrumento común para una perspectiva política «posible». Con sociedad open source se hace referencia a una sociedad cuyo código sea abierto y en cuya mejora pueda participar todo el mundo libremente. Puesta en estos términos no se puede no compartirla. Pero sorprende la ligereza con la que se intercambia un término procedente de un recorrido particular, técnico e informático, para generalizarlo, aplicándolo a teorías filosóficas, económicas, sociológicas, sin considerar las posibilidades de modificar aquel concepto o cuando menos utilizarlo con las debidas aclaraciones.

Open source, de hecho, en el campo del software en el que ha surgido el término, significa también competencia, competición por hacerse con los mejores cerebros al menor precio, financiaciones de capitales de riesgo, operaciones millonarias de compra y venta, un gran negocio orientado a una forma más «democrática» y suave de capitalismo. Esta dinámica intenta no ya someter la fuerza de trabajo, sino cooptarla en la realización de la misión de la empresa, que se identifica cada vez más con la realización de los propios deseos particulares[10].

Entre las muchas sociedades que hoy en día se han sumado a esta tendencia para sacar el máximo partido posible también está Google que, por otra parte, es bueno por definición: por ello, *don't be evil,* usa el software libre, es gratuito, es mejor que el propietario, los desarrolladores están orgullosos de colaborar. La visita a Googleplex ha mostrado claramente como en Mountain View esta estrategia de penetración en las vidas de las personas se ha pulido hasta el máximo grado: trabajadores gratificados, animados a ser creativos y felices producen más y mejor que los trabajadores frustrados y oprimidos.

10 En un contexto atrasado como el italiano, en el que las empresas, incluso en el sector de las tecnologías de la información, continúan aplicando unas obsoletas lógicas fordistas de producción en masa, sin valorar mínimamente las potencialidades de los individuos, podría parecer despreciativo el ataque al capitalismo de la abundancia de Google. Pero como no todo el mundo cree en el desarrollo sostenible y en el consumismo responsable promovidos por una hábil campaña de propaganda tercermundista, entonces no podemos avalar ninguna forma de explotación de los individuos, ni siquiera la de Googleplex, sutil e incluso agradable para los trabajadores convertidos ya en creativos entusiastas. No se trata de rigidez dogmática, sino de una mínima ética común (*common decency*, diría Orwell): el objetivo último no es la plena realización de la persona, sino el despliegue sin fin del capital, no importa si suave o ferozmente represivo. Richard Sennett lo ha demostrado con lucidez y un amplio examen histórico en *La corrosión del carácter: las consecuencias personales del trabajo en el nuevo capitalismo*, Anagrama, Barcelona, 2000. Por otro lado, el consumismo extremo del capitalismo de la abundancia podría constituir el primer paso hacia un nuevo fascismo, como muestra la lúcida visión de J. G. Ballard en *Milenio negro* (Minotauro, Barcelona, 2004) y en *Kingdome Come*.

Seducir a los hackers: autonomía, dinero fácil e instrumentos gratuitos

La explotación del open source por parte de Google alcanza su cima alrededor del 2005, cuando su imagen se ve empañada por las acciones de la competencia y por vicisitudes legales no del todo claras[11].

Por mucho que el proyecto estuviese enraizado de forma notoria en la cultura informática y en la práctica de la excelencia académica, no era suficiente usar el sistema operativo GNU/Linux para que el datacenter de Google funcionara: hacía falta explicitar la confianza hacia el open source con una iniciativa fuerte, que reclamara la atención en el magma de las redes de producción libre.

Ya no bastaba ofrecer el soporte para el idioma h4x0r —el lenguaje de los «verdaderos» hackers— (o para el klingoniano de Star Trek) a fin de hacerse con las simpatías de los desarrolladores. Además, la actitud elitista de los supercocos universitarios empezaba a cansar a los inversores.

Arrogancia y culto meritocrático de derivación académica, aunque siempre con óptimos resultados, tienen poca respuesta por parte de los inversores que quieren dividendos sustanciales. Era inevitable que se cerrase la fase en la que los dos podían permitirse iniciar su cotización en bolsa con la pretensión de reunir, tras la primera oferta de acciones, la cantidad exacta de 2.718.281.828 dólares, un número que repite la constante matemática «e» (la base de la función logaritmo natural); o también salidas alegres como el anuncio en agosto de 2005, cuando declaran que iban a vender 14.159.265 acciones para recoger 4.000 millones de dólares líquidos, sin dar explicaciones a los inversores del uso que iban a hacer de aquel dinero.

Para sostener concretamente el propio deseo de invertir en investigación, para demostrar cómo mediante una estrategia semejante no solo se puede competir, sino incluso ser los mejores en el mercado, hacía falta dar un paso estratégico dirigido no tanto a los usuarios «normales», sino más bien a los jóvenes cerebros, al futuro, a la innovación. Traducido operativamente, esto significa crear comunidades, ofrecer instrumentos de desarrollo, sellar acuerdos con otras sociedades del sector. O sea, cortejar el mundo del open source.

Google comienza a invertir en la creación de comunidades en octubre 2005, cuando financia con 350.000 dólares la Oregon State University y la Portland State University. Sucesivamente se pone en marcha con gran pompa el programa de junio de 2005: *Summer of Code*, es decir, «el verano del código», promovido directamente en una página del sitio de Google y localizable en http://code.google.com/summerofcode05.html.

El estilo comunicativo queda bien claro: ofrecer mejores oportunidades. Todo programador que ya hubiese creado un proyecto de open source nuevo o hubiese aportado una mejora digna de mención a un proyecto ya existente durante el verano, iba a recibir un premio de 4.500 dólares. La operación estaba evidentemente enfocada a presentarse como una explícita declaración de amor hacia el open source, o sea, subrayar que el open source era el terreno estratégico sobre el que cultivar la innovación. En segundo lugar, miraba de atraer las simpatías de los jóvenes desarrolladores mediante una operación de apoyo financiero concreto a su trabajo. Así pues, intentaba crear una verdadera y propia comunidad de estilo «abierto», por lo menos en lo referente a los patrocinadores.

Los programadores premiados, en su mayoría estudiantes, fueron más de cuatrocientos; por lo general, aportaron modificaciones e introdujeron novedades sobre proyectos ya existentes, en lugar de dar a conocer sus nuevos softwares, añadiendo características a programas como Apache, Fedora, Gaim, el mismo Google, Inkscape, Jabber, KDE, Mozilla, OpenOffice.org, Pitón, Samba, Gnome, Mono, Ubuntu. Una buena ganancia para todos, especialmente para las socieda-

11 Véase capítulo I.

des que están detrás de estos proyectos: entre las principales, recordemos IBM, RedHat, LinSpire, Novell, Mozilla.com, Sun, HP, Ubuntu[12].

Algunos de estos proyectos, junto con los desarrollados en el notorio 20% de tiempo libre por los propios empleados de Google, han permitido alcanzar el segundo objetivo en el recorrido de acercamiento al open source: ofrecer instrumentos de desarrollo. Ya desde 2002 Google ofrecía herramientas de desarrollo que podían descargarse gratuitamente del sitio code.google.com. Ahora esta página contiene los proyectos propietarios creados por los equipos de desarrollo de Google junto con los proyectos ganadores de Summer of Code que, de alguna manera, están relacionados con sus servicios.

La sección «Code» del sitio propone algunos proyectos dedicados a los creadores de software en los más variados lenguajes de programación (Java, C++, Python, etc.). Facilitar herramientas de desarrollo es un elemento esencial para cualquiera que desee permitir la creación de software y comunidad, porque se invierte precisamente sobre los medios de trabajo necesarios para su creación. Los proyectos ideados por los programadores de Google como herramientas de desarrollo se conocen con el nombre de Google API, librerías propietarias que permiten la interfaz y la utilización de los principales servicios del coloso de Mountain View.

Una librería es un conjunto de funciones compartidas: se trata de porciones de código completadas que facilitan herramientas a otros programas que necesiten de funciones simplificadas. Un ejemplo muy claro lo constituyen las librerías gráficas GTK, QT y FLTK, que implementan los elementos estándares de una aplicación visual (botones, menú, iconos...), simplificando el trabajo de los programadores: estos últimos, apoyándose en la librería escogida, escribirán solamente las funciones reales del programa. De hecho, serán las

12 *Summer of Code*: http://code.google.com/soc/. El evento, que se repite cada año, tiene el éxito garantizado.

propias librerías las que dibujen los botones, gestionen los clics del ratón, diseñen las sombras y lleven a cabo buena parte de lo que estamos acostumbrados a ver como usuarios. El tiempo y las capacidades de los programadores así se optimizan; dado que quien escribe código difícilmente será un entusiasta de la creación de botoncitos para las propias aplicaciones, las librerías gráficas adquieren un papel esencial de enlace entre los diferentes proyectos. Por un lado, las aplicaciones aparecerán relativamente homogéneas desde el punto de vista gráfico; por el otro, los programadores podrán centrarse en su trabajo, sin perder demasiado tiempo en la implementación de las interfaces.

Existen comunidades de desarrollo que se ocupan de las librerías para ofrecer herramientas genéricas y transversales para la solución de problemas complejos (conexiones de red, comunicación entre las aplicaciones, gestión del texto, imágenes, compresiones). Exactamente de la misma forma en la que se escribe un software con el objetivo de alcanzar el máximo de usuarios posibles, así se confecciona una librería para que llegue al mayor número de desarrolladores.

Las librerías, por lo tanto, permiten a los programadores la creación de sus propios softwares partiendo de un set de elementos compartidos, verdaderos y auténticos estándares de facto de la programación. Apoyarse en las librerías significa implementar el trabajo valiéndose de una base de partida muy amplia y compleja, que utiliza de forma óptima el código ya disponible y estratifica las competencias. Las librerías tienen, pues, un valor estratégico tanto en la dinámica de la cooperación espontánea del free software como en la economía relacional del open source.

Las librerías de Google, o sea Google API, están publicadas bajo licencias propietarias, es decir, ocultan el mecanismo de su funcionamiento al programador. Pero eso no es todo: incorporan también un especial dispositivo de control, el desarrollador que descarga gratuitamente las librerías tiene que autenticarse mediante un código. Este sistema le permite a

Google delinear de forma invasiva todos los movimientos y las modificaciones que proceden del uso de sus API. Los programadores que usan estas librerías tienen la oportunidad de insertar la búsqueda de Google en su propio sitio o de conocer en tiempo real el propio PageRank(™). Además, pueden realizar softwares capaces de gestionar campañas publicitarias mediante AdWords, generar mapas dinámicos de sus datos con la interfaz de Google Maps o incluso implementar un client VoIP para la telefonía online compatible con Gtalk. En suma, pueden desarrollar los servicios de Google como mejor crean, en el lenguaje de programación que prefieran, bajo la atenta supervisión de Mountain View.

La enorme difusión de los servicios de Google se conjuga con la posibilidad de personalización en los mínimos detalles: de hecho, mediante la escritura de los útiles documentos XML[13], es posible crear «puentes» de conexión entre los distintos servicios de Google; por ejemplo, programando pieza por pieza la *homepage* de Google como si fuese una verdadera aplicación, volviéndola, así, totalmente adecuada a las propias necesidades. Algo muy parecido puede hacerse con Google Earth: es posible construir navegaciones especiales en 3D sobre las fotos de satélite, subrayando gráficamente en los ordenadores de los usuarios zonas geográficas, edificios, datos climáticos, etc.

Todas esas herramientas predispuestas para quien sabe escribir código (al menos en un lenguaje) son esenciales para

13 XML, *eXtensible Markup Language*, es un lenguaje extensible realizado para poder utilizar de forma simple los documentos estructurados, pensado para la Web y para superar las limitaciones de HTML (*Hyper Text Markup Language*), pero con posibilidad de ser utilizado también en ámbitos diferentes. Desarrollado a partir de W3C, World Wide Web Consortium, XML es un subconjunto de SGML (*Standard Generalized Markup Language*), un estándar internacional que define las reglas para escribir *markup language*; intencionalmente no incluye algunas funcionalidades complejas de SGML que difícilmente pueden ser implementadas en la Web. El primer esbozo de XML se remonta a noviembre de 1996, el actual es consultable en la dirección http://www.w3.org/TR/1998/REC-xml-19980210.

hallar nuevas combinaciones o simplemente para usar lo que Google hace público, cuando menos en parte, dentro de sus propias aplicaciones[14]. Incluso existe un portal, googleearthhacks.com, en el que se pueden encontrar muchas trampas y *hacks* para utilizar de las formas más impensables este servicio, cruzando los mapas del satélite con cualquier otra base de datos.

Todas las posibilidades que las API de Google facilitan implican el respeto de dos reglas especiales: el registro y la licencia. Para activar funciones de Google API es preciso solicitar una clave, es decir, un código de acceso, y comunicar exactamente dónde se las quiere emplear. Las API se activan solo después de haberse registrado. El segundo punto es la licencia. Ya que estas API no tienen una licencia copyleft, se pueden utilizar solo con ciertas limitaciones: por ejemplo, es preciso tener una cuenta Google para que la acumulación de informaciones no se pare nunca; además, los mapas son de propiedad exclusiva de Google (o de terceras partes) y no pueden ser modificados bajo ningún concepto. Naturalmente, para usos comerciales es necesario estipular un contrato. El código de activación permite que Google mantenga siempre el control total sobre los nuevos programas generados sobre las API: puede bloquear a su gusto las aplicaciones o simplemente controlar tanto la manera en la que se accede a los servicios, como el uso que se hace de ellos. Todo esto es posible porque el código fuente no es público y libre y, por lo tanto, resulta imposible comprender el funcionamiento interno de las librerías.

Además de dejar desarrollar gratuitamente y de monitorear el desarrollo de los propios servicios, otra razón por la que Google está creando comunidad con esta extraña fórmu-

14 Sorprende el hecho de que el motor de búsqueda, desde el punto de vista de la interacción con las API, esté limitado a mil solicitudes diarias. Esto significa que los softwares que se basan en las API podrán realizar como mucho mil búsquedas diarias. Probablemente las limitaciones disminuirán con el tiempo, porque la capacidad de Google crece sin descanso; sin embargo, Google impone las limitaciones sin ofrecer ninguna explicación.

la que podríamos definir como pseudo-open es la de conseguir una base de datos sobre la que realizar investigación y venta de estadísticas.

Hospedar gratuitamente los proyectos de los programadores individuales significa conseguir su confianza. Permitir a cualquiera realizar búsquedas en la base de datos de los proyectos hospedados activa una sólida cadena de usuarios. Una incubadora así de jóvenes talentos garantiza, además, la disponibilidad de material humano altamente motivado y cuya formación, es decir el coste principal en el sector de las tecnologías de la información, ya se ha cumplido de forma autónoma y en conformidad total con el estilo de la empresa.

La oferta de herramientas de desarrollo es un mecanismo de *talent-scouting* conocido desde hace tiempo y en particular es el caballo de batalla de algunas sólidas empresas como la Va Software Corporation, que pone a disposición gratuita para el mundo del open source ordenadores extremadamente potentes y banda ancha, espacio de disco y asistencia no al alcance de cualquiera. Dos paraísos digitales pueden presumir de una fama mundial y de un número de proyectos hospedados superior a toda la competencia: SourceForge.net y freshmeat.net, ambos de propiedad de la Va Software. Estos dos portales tienen una resonancia tal que incluso los proyectos más pequeños, tras aparecer en las primeras páginas, llegan tranquilamente a contar con centenares de visitas diarias. Todos los proyectos en el seno de code.google.com tienen su correspondiente página en freshmeat.net y/o en SourceForge.net.

De esta forma, las aplicaciones pueden a la vez gozar de la visibilidad de Google y de todos los servicios puestos a disposición por el coloso Va Software: mailing lists, ordenadores dedicados a la solución de errores de programación (*debug*), a sistemas de control de las versiones, de las revisiones y de los cambios del código (por ejemplo, *CVS, Concurrent Versioning System*), fórums de discusión, etc.

Es fácil imaginar como, con una base de datos utilizada gratuitamente por millares de *coders*, la Va Software puede

garantizar un óptimo servicio de *business to business* a las empresas relacionadas en mayor o menor medida con el open source y a otras sin ninguna relación con éste. Un *data mining*[15] de especial interés para negocios millonarios. Entre los patrocinadores y los anunciantes de SourceForge.net encontramos a Red Hat, Microsoft y muchos más.

Hay muchas formas de poner en contacto los desarrolladores con las empresas open source. En Italia, SUN Microsystem ofrece la posibilidad de publicar el propio currículum sobre un mapa de Google (utilizando las Google API), mediante el portal javaopenbusiness.it. Los mismos desarrolladores son los que indican su propio perfil, creando así un mapa de las competencias open source en Italia a través de las herramientas que SUN y Google han puesto a disposición.

Google puede contar pues con la implementación prácticamente gratuita de los propios productos por parte de centenares de usuarios; a todo esto se añade la inversión interesada en competiciones como *Summer of Code*, festivales dedicados a la promoción y desarrollo de los propios servicios y, por último, aunque no menos importante, sistemas de reclutamiento excepcionalmente dinámicos. Entre éstos, también se sitúa el vídeo-reclutamiento, directamente en las páginas de vídeo.google.com, con entrevistas a empleados entusiastas y a Sergey Brin en persona, todos de acuerdo en ilustrar los privilegios de trabajar en Mountain View[16].

15 «El *data mining* (minería de datos), es el conjunto de técnicas y tecnologías que permiten explorar grandes bases de datos, de manera automática o semiautomática, con el objetivo de encontrar patrones repetitivos, tendencias o reglas que expliquen el comportamiento de los datos en un determinado contexto». Fuente: http://www.sinnexus.com/business_intelligence/datamining.aspx). [N. del E.]

16 El vídeo está disponible en la dirección http://www.youtube.com/watch?v=w887NIa_V9w&feature=channel.

Ambientes híbridos entre universidad y empresa

Puestas estas premisas, el acercamiento de Google al open source parece bastante estratégico e interesado, por mucho que se haya originado gracias a un sentir común respecto a las dinámicas cooperativas típicas de las comunidades de desarrollo free software, surgidas en el humus académico. La estrategia de acumulación evidenciada con anterioridad también avanza en este ámbito: de hecho Google se comporta como una especie de agujero negro que usa códigos abiertos, o incluso favorece la redacción de esos códigos y los atrae, para luego introducirlos en su propio circuito. Por ejemplo, ninguna de las modificaciones que los programadores de Google han aportado a las herramientas abiertas usadas se ha hecho pública jamás. Es de destacar su Google Web Server (GWS), una versión modificada de una versión de Apache y el servidor web open source más difundido en la Red. Esto significa, sin duda, explotar las potencialidades y las realizaciones del método de desarrollo abierto, sin compartir, sin embargo, las propias implementaciones y mejoras.

Un factor de importancia primaria a propósito de las relaciones con el mundo open source es que Google nace en Stanford, una universidad conocida por su capacidad de generar *start-up* agresivas y competitivas basándose en investigaciones de perfil elevado. Por mucho que Stanford fuera, y siga siendo, un ambiente favorable al desarrollo de proyectos open source, su unión de doble filo con el capital de riesgo hace que sea difícil, más bien imposible, seguir adelante en la vía de la excelencia académica una vez se salga del campus.

Será necesario un breve esbozo sobre la investigación académica americana para comprender los orígenes de Google, entre el open source y la investigación orientada al provecho. A nivel más general hay que subrayar el carácter centralizador de la universidad estadounidense a propósito de la creación intelectual: todos los proyectos desarrollados en campo académico tienden a ser copyright de la universidad que ha hospedado el equipo de investigación. Stanford no es una excepción; por otra parte, en los EE.UU. las instituciones académicas están relacionadas históricamente con el mundo de los negocios y, a menudo, son auténticas empresas. Las patentes universitarias sobre las invenciones de los investigadores generan royalties muy sustanciales; confieren, además, prestigio a los centros de investigación y a los estudiantes/investigadores/empresarios.

Las universidades son ambientes híbridos, entre lo público y lo privado. En los EE.UU. hasta 2002, cuando menos en teoría, los lugares de investigación públicos no podían patentar sus invenciones; lo mismo vale para los laboratorios privados pero financiados con fondos públicos (por lo tanto, a menudo también las universidades). El pago de aranceles obstaculiza la libre circulación de los saberes en la investigación científica, la posibilidad de reproducir, comprobar o falsificar los resultados experimentales. Y esto era así por el *Experimental Use Defense*, «protección del uso experimental», una norma que autorizaba el uso gratuito de tecnologías patentadas en el ámbito de la investigación, introducida en 1813, y abolida en 2002 con una sentencia a favor del investigador John Madey. Madey demandó a la Duke University, para la que trabajaba, porque usaba una maquinaria patentada por él para investigaciones láser sobre electrones libres. La Corte consideró que el *Experimental Use Defense* había surgido para proteger al científico dedicado a la investigación libre y desinteresada, pero evidentemente en las universidades esta actividad ya no es tan inocente e, incluso en el caso en el que no sea directamente comercial, puede ser considerado como un «negocio lícito» *(legitimate business*), ya que conlleva financiaciones y precisa de fuerza de trabajo y de personal en formación (los estudiantes). Cae así toda distinción entre la investigación privada y la pública[17].

17 Noticias comentadas por Laser en http://www.e-laser.org/htm/new.

Naturalmente, todos los proyectos nacidos en Stanford son patentados por parte de la universidad, y esta mezcla entre incentivo a los proyectos open source, por un lado, y patentabilidad salvaje, por el otro, no beneficia al ideal, ni mucho menos a la práctica de la «investigación» en sí, tan ostentada como motivo de orgullo y de fuerza de Google

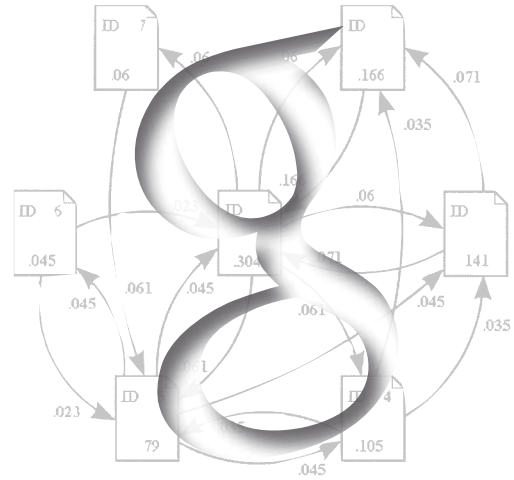
La cuestión de la patente se hace aún más interesante si recordamos que el éxito de Google se basa en un algoritmo ideado por Larry Page, en colaboración con Sergey Brin, cuando aún eran investigadores en la Facultad de Ciencias Informáticas en Stanford. El algoritmo que ha revolucionado la indexación de la Red es por lo tanto propiedad de Stanford, sometido a patente regular. Vamos a descubrir al detalle cómo funciona este prodigio, cómo consigue ofrecer resultados en tiempos más rápidos que cualquier otro de la competencia, pudiendo dar casi de verdad a cada usuario «lo que desea».

asp?idNews=223 y http://www.e-laser.org/htm/new.asp?idNews=151; y en el libro del Gruppo Laser, *ob. cit.*, http://infofreeflow.noblogs.org/post/2008/02/10/il-sapere-liberato/; fuentes directas sobre el caso Madey: http://www.law.nyu.edu/journals/lawreview/issues/vol792004/number4/index.htm#top.

4

Algoritmos,

¡qué pasión!



El crecimiento vertiginoso de Google no ha mermado su fama de motor rápido y eficiente, fiable y completo: todos hemos oído decir que «¡si no está en Google, no existe!» y que «Google es más veloz». En la base de este éxito, además de los elementos que hemos ido analizando hasta ahora, está el algoritmo de PageRank(™), citado al comienzo de este texto, que guía el spider de Google en el descubrimiento de las redes. Veamos en detalle de qué se trata y como funciona.

Algoritmos y vida real

Un algoritmo[1] es un método resolutivo aplicado a un problema, un procedimiento compuesto por pasos simples a ejecutar en secuencia para obtener un determinado resultado. A un algoritmo que favorece la solución de un problema se le

1 «Algoritmo: conjunto de reglas o directivas aptas para ofrecer una respuesta específica a uno o más datos de input». Para una primera introducción, http://es.wikipedia.org/wiki/Algoritmo. El término algoritmo deriva del nombre de «al-Khwarizmi», importante matemático árabe del siglo IX. Muhammad ibn Musa al-Khwarizmi introdujo el uso de los números árabes en la matemática: su obra *(Libro) de al-Khwarizmi sobre los números indios* fue traducida en latín como *Algorismi de numero Indorum*. Sobre la base de este texto toda Europa aprendió a utilizar el

denomina «correcto», y si la solución se obtiene en un tiempo breve se le llama «eficaz». Existen muchos tipos distintos de algoritmos, empleados en campos muy diferentes de las ciencias; pero no se trata de procedimientos abstractos propios de una exigua minoría de estudiosos, sino de prácticas que influencian nuestras vidas cotidianas mucho más de lo que nos parece a primera vista.

Por ejemplo, las técnicas para grabar un programa televisivo utilizan algoritmos, y también los métodos para ordenar una baraja de naipes o para planificar las paradas de un viaje particularmente largo. En un tiempo relativamente previsible, realizando una serie de pasos sencillos y repetibles de manera idéntica, elegimos más o menos implícitamente los algoritmos adecuados a la solución que estamos buscando. «Sencillo» significa sobre todo unívoco, evidente para quien aplicará el algoritmo. En este sentido, una receta es un algoritmo: «Ponga a hervir tres litros de agua en una olla, añada sal y eche quinientos gramos de arroz, cuélelo pasados doce minutos, añada especias a voluntad», es una descripción de pasos sencillos y unívocos, si el destinatario de la receta está en condiciones de discernir pasos como «poner sal» o «añadir especias a voluntad».

Los algoritmos no son necesariamente métodos para hallar una solución en el menor tiempo posible. De hecho, hay algunos que se ocupan de obtener soluciones aceptables sin preocuparse del factor tiempo; otros permiten conseguir un resultado en el menor número de pasos, o su prioridad es el ahorro de recursos[2].

sistema de numeración decimal posicional hoy todavía en vigor; la numeración romana fue progresivamente abandonada. Los procedimientos que permitían efectuar cálculos en numeración decimal fueron denominados «algorismi» o «algoritmi», y más tarde el mismo término fue aplicado en general a los procedimientos de cálculo necesarios para obtener un determinado resultado.

2 El mejor método para llegar a París puede ser salir con un vuelo directo desde el aeropuerto más cercano a nuestra ciudad, o coger el primer vuelo de salida del mismo aeropuerto, bajar a la estación de llegada, volver

Nos importa subrayar aquí, más allá de cualquiera profundización especializada, la naturaleza práctica, aplicativa de los algoritmos; nos afectan a todos pues son prácticas concretas para alcanzar un determinado objetivo. En el campo informático se utilizan para solucionar problemas recurrentes en la programación del software, en los proyectos de redes y en la construcción de aparatos de hardware. A lo largo de los últimos años, sobre todo por la creciente importancia de los modelos reticulares de análisis e interpretación de la realidad, muchos investigadores han enfocado sus estudios hacia las metodologías de construcción y de recorrido de las redes y de los datos que constituyen su materia viva. La economía de la búsqueda de la que habla John Battelle[3] se cumple gracias al perfeccionamiento de algoritmos para la búsqueda de informaciones, estudiados para aumentar la posibilidad de encontrar y compartir datos de forma cada vez más eficiente, veloz, fiable y segura. El caso más conocido por el gran público es el fenómeno del *peer-to-peer*: en lugar de crear enormes bancos de datos a los que acceder para encontrar vídeos, audios, textos, software y todo tipo de informaciones, se desarrollan continuamente algoritmos cada vez más optimizados para facilitar la creación de redes altamente descentralizadas, en las que cada usuario puede ponerse en contacto directamente con otros usuarios y llevar a cabo provechosos intercambios[4].

a coger el primer avión, bajar y nuevamente volver a partir, continuando hasta llegar a París. Es bastante cierto que los dos métodos nos permitirán llegar a la destinación: con el primer método llegaremos a la meta en el menor tiempo posible y probablemente con el mayor ahorro de dinero; el segundo, en cambio, nos permitirá apreciar los principales aeropuertos de las capitales europeas en las diversas épocas del año. Algoritmos diferentes describen las distintas posibilidades.

3 Véase el capitulo II.

4 Generalmente, por *peer-to-peer* (o P2P) se entiende una red de ordenadores o cualquier red que no posea *clients* o *servers* fijos, sino un número de nodos equivalentes (*peer* significa «pares») que hacen funciones bien sea de *client* bien sea de *server* hacia otros nodos de la red. Este modelo de

La estrategia de la objetividad

El crecimiento exponencial de la calidad y cantidad de ancho de banda de nuestros ordenadores, junto con la constante disminución de los costes, nos ha permitido navegar en Internet mejor, durante más tiempo y más rápidamente. Hace nada más veinte años los módems con pocos baudes (número de símbolos transmitidos por segundo) eran un lujo para pocos, mientras que ahora también en Italia la fibra óptica, por la que viajan millones de bytes por segundo, es una tecnología accesible.

Hace diez años hacían falta notables conocimientos informáticos para crear contenidos aptos para las redes digitales; ahora, en cambio, la mayor facilidad de publicación de contenidos en la Web, la omnipresencia del correo electrónico, las mejoras de los sistemas de escritura colectiva online, como los blogs, wikis, portales, mailing lists; y la bajada, al mismo tiempo, de los costes de grabación y mantenimiento de los dominios y de los espacios de Internet favorecen la transformación progresiva de los usuarios desde simples usuarios de informaciones, puestas a disposición por especialistas de las tecnologías de la información, a creadores de informaciones.

La mejora de la conectividad se incrementa por tanto a la par que el crecimiento exponencial de los datos volcados en la Red y, como ya hemos anotado, implica la urgente necesidad de cada vez mejores herramientas de búsqueda. No es de extrañar que la urgencia de servicios de búsqueda atraiga el interés de sociólogos, informáticos, economistas, diseñadores y estudiosos de la comunicación en general. Por otro lado,

red es la antítesis de la arquitectura client-server, donde existe una relación jerárquica entre el proveedor del servicio (*server*) y el receptor (*client*). Por medio de esta configuración distribuida, cualquier nodo está en condición de poner en marcha y completar una transacción. Los nodos equivalentes pueden diferir de la configuración local, en velocidad de elaboración, en amplitud de banda y cantidad de datos memorizados. El ejemplo clásico de P2P es la red para la puesta en común de *files* (fuente: http://www.inventati.org/inventa/doku.php?id=glossario_no-copyright_ecc.).

el diluvio informacional de las redes globales no es una simple «puesta en red» de las sociedades tal y como las conocemos, sino un fenómeno extremadamente complejo, que exige interpretaciones no simplistas. Creemos, por tanto, que dicho compromiso teórico y práctico no se puede delegar a los especialistas, sino que debe ser el resultado de una elaboración colectiva.

De hecho, si por un lado la construcción de redes autogestionadas puede ser una oportunidad para ampliar y poner en contacto entre sí zonas autónomas, por el otro el control social encuentra en las tecnologías de la información un formidable instrumento de represión.

La realización de este segundo escenario, del que el caso Echelon[5] es tan solo la manifestación más clamorosa, parece ciertamente cada vez más probable, a la luz del constante aumento del número de individuos que producen informaciones, y entra en abierta contradicción con la disminución continua de los proveedores de herramientas de búsqueda. El acceso a las informaciones producidas por un número cada vez mayor de individuos está dirigido por un puñado de monopolistas que reducen una delicada cuestión social y política a una competición de marketing sin ahorro de golpes, en la que la adopción de un mejor algoritmo resulta ser la clave de la victoria.

Un algoritmo de búsqueda es un instrumento técnico que activa un mecanismo de marketing extremadamente sutil: el usuario se fía de que los resultados no sean filtrados y correspondan a las preferencias de navegación que genera la comunidad de usuarios. En síntesis, se difunde un mecanismo de confianza en la objetividad de la técnica (específicamente, el

5 Echelon es el nombre de un sistema de control y espionaje electrónico global puesto en marcha por el Gobierno de los Estados Unidos. Su justificación oficial es hoy la lucha en contra del terrorismo: comunicaciones telefónicas, mails y toda comunicación digital se intercepta y analiza con fines políticos y económicos. Para un análisis en profundidad véase: Duncan Campbell, *Il mondo sotto sorveglianza-Echelon e lo spionaggio elettronico globale*, Eléuthera, Milán, 2003, http://home.hiwaay.net~pspoole/echres.html recoge muchísimos recursos disponibles.

procedimiento algorítmico que genera el resultado de la interrogación), que se considera «buena» en tanto que no influenciada por las idiosincrasias y preferencias de individuos humanos. Las máquinas «buenas», hijas de una ciencia «objetiva» y de una investigación «desinteresada», no manipularán los resultados, no nos contarán mentiras porque no pueden mentir y, de todas formas, no tendrían ningún interés en hacerlo. Bien distinta es la realidad, y esta creencia se revela una hipótesis demagógica, detrás de la cual las máquinas del marketing y del control acumulan fabulosas ganancias.

El caso de Google es el ejemplo meridiano de esta «estrategia de la objetividad» asociada a la técnica: el motor de búsqueda «bueno por definición» explota y graba por entero y de forma continua los comportamientos de los usuarios que utilizan sus servicios, con el fin de identificar sus costumbres e insertar en sus actividades (navegación, correo, gestión de files, etc.) anuncios personalizados, contextuales, ligeros, omnipresentes y, posiblemente, en condición de generar *feedback,* de modo que los usuarios ofrezcan informaciones útiles para los vendedores y hasta lleguen a mejorar por sí mismos «las sugerencias publicitarias», expresando sus preferencias. La continua petición de la opinión de los usuarios, además de halagar a las personas que se sienten participes de una extensa «democracia electrónica», es de hecho la forma más simple y eficaz para conseguir preciosas informaciones comerciales acerca de los gustos de los consumidores. Son las preferencias (y la inconsciencia) de los usuarios las que hacen triunfar un motor de búsqueda sobre los demás, pues un sitio muy visitado puede modificar sus contenidos sobre la base de «sugerencias» comerciales y, consecuentemente, activar hábiles movimientos económicos.

Desde un punto de vista exquisitamente informático, a los motores de búsqueda les pertenece la gestión de cuatro elementos: la búsqueda de datos en la Red (spider), la memorización de las informaciones en archivos (bases de datos), un algoritmo válido para ordenar los datos según las búsquedas

realizadas (interrogación) y, por ultimo, el desarrollo de una interfaz capaz de satisfacer al usuario; cada uno de los tres primeros aspectos está cuidado por un tipo específico de algoritmo: búsqueda, memorización/archivación, interrogación.

La potencia de Google, como la de Yahoo! y otros gigantes de la búsqueda en la Red, se basa por tanto en:

1. un *spider*, o sea, un software para extraer contenidos de las redes;
2. hard-disks de enorme capacidad para memorizar los datos en soportes fiables para evitar toda pérdida accidental;
3. un sistema rápido para encontrar (y ordenar) los resultados de una interrogación sobre la base del valor del *ranking* de las páginas;
4. una interfaz usuario vía web (y no solo: Google Desktop y Google Earth, por ejemplo, son programas que hay que instalar en la máquina del usuario) para responder a las peticiones acerca de estas informaciones.

Spiders, bases de datos y búsquedas

El *spider* es una aplicación que, en la mayoría de los casos, se desarrolla en los laboratorios de investigación de los propios motores de búsqueda. Su finalidad es navegar saltando de un link a otro en las páginas web recogiendo informaciones: formatos de documentos, palabras clave, autores de las páginas, links ulteriores, etc. Al final de sus exploraciones el software-spider lo entrega todo a la base de datos que archivará las informaciones. Además, el spider debe ocuparse de captar las variaciones de cada sitio y entonces programar visitas sucesivas para almacenar nuevos datos. Concretamente, el spider de Google gestiona dos tipologías de chequeo de los sitios: una mensual en profundidad, *deep-crawl,* y una diaria de puesta al día, *fresh-crawl.* De esta forma, el spider actualiza la base de datos de Google, recogiendo los cambios en las redes. Después de una exploración exhaustiva, Google

necesita algunos días para actualizar las diferentes indexaciones y difundir los nuevos resultados en todos los *datacenters*. Este lapso de tiempo es conocido como *Google dance* (danza de Google): los resultados de las búsquedas pueden diferir sensiblemente, pues se refieren a índices diferentes. A partir de 2003 Google modificó sus metodologías de catalogación y actualización, limitando drásticamente los efectos de la «danza» y diluyéndolos en el tiempo: de hecho, ahora los resultados de las búsquedas varían de manera dinámica y continua sin ningún desbarajuste periódico. En realidad, los resultados de las búsquedas difieren también sobre la base de las anteriores navegaciones de los usuarios, que se archivan y utilizan para «mejorar», en el sentido de «simplificar», la recogida de las informaciones[6].

La secuencia de elecciones que la aplicación lleva a cabo para indexar un sitio es la verdadera potencia del algoritmo de Google. Mientras que el algoritmo de base de PageRank(™) está depositado bajo patente de Stanford y es, por tanto, público, estos ulteriores pasos algorítmicos no son publicados ni por Google ni por ninguno de los motores de búsqueda actualmente en uso; ni tampoco son públicos los procesos de recuperación de la base de datos.

En el ámbito informático, una base de datos (*database*) es, esencialmente, un archivo digital; en su forma más sencilla —y actualmente más difusa— se puede representar bajo la forma de una o más tablas en relación entre sí que presentan valores de entrada y de salida: se habla entonces de database relacional. Como todo archivo, una base de datos está organizada según reglas precisas de almacenaje, extracción y mejora continua de los propios datos (recuperación de datos corruptos, corrección de entradas duplicadas, constante reingenierización de los procesos de captación de los datos, etc.).

Los informáticos estudian desde hace decenas de años las metodologías de búsqueda, introducción y mejora de los datos en la database, experimentando lenguajes de programación y diferentes acercamientos (jerárquico, reticular, relacional, objetual, etc.). La fase del proyecto de una base de datos es un componente crucial del proceso de desarrollo de un sistema informativo complejo como el de Google, pues de ella depende en esencia su funcionalidad. Para conseguir una rápida extracción de los datos y, en general, una gestión eficiente es fundamental la correcta individuación de las finalidades de la database y —en el caso de las bases de datos relacionales— de las tablas, a definir a través de sus campos y de las relaciones que las unen. Naturalmente es necesario adoptar inevitables aproximaciones en los pasos entre las lenguas naturales, analógicas, y los datos introducidos, digitales. El aspecto negativo es el secretismo de estas metodologías: como sucede en todos los proyectos propietarios de desarrollo, a diferencia de aquellos libres, es harto difícil saber qué instrumentos y qué algoritmos han sido utilizados.

A través de los textos redactados por los centros de investigación y por las universidades se pueden encontrar las escasas informaciones hechas públicas por los proyectos propietarios. En estos textos se encuentran informaciones útiles para comprender la estructura de los ordenadores y la gestión de los datos por parte de los motores de búsqueda. Para hacernos una idea de la potencia de cálculo actualmente disponible, se describen ordenadores capaces de convertir direcciones de Internet en secuencias unívocas de bytes útiles como índices para las bases de datos en 0,5 microsegundos, y capaces de ejecutar 9.000 spiders simultáneamente; concretando, se trata de sistemas capaces de analizar y almacenar cerca de 50 millones de nuevas páginas al día[7].

El último elemento algorítmico escondido detrás de la «sencilla» fachada de Google es el dispositivo de búsqueda, o sea, ese sistema que, dada una interrogación, es capaz de en-

6 Véase capítulo V.

7 Véase por ejemplo la documentación publicada por el IBM Almaden Research Center: http://www.almaden.ibm.com/webfountain/publications/.

contrar los resultados más adecuados, ordenarlos por importancia y ranking, y enviarlos a la interfaz.

Algunas universidades y laboratorios han decidido publicar sus investigaciones en tal sentido, particularmente las soluciones halladas y los diferentes métodos utilizados para optimizar la velocidad de acceso a las informaciones, la complejidad de la ordenación y la selección de los parámetros de inputs más interesantes.

Los motores de búsqueda, de hecho, deben ser capaces de ofrecer resultados óptimos casi instantáneamente, ofreciendo al mismo tiempo un abanico de posibilidades de elección lo más amplio posible. Google representa sin duda el estado del arte de los motores de búsqueda: tales extraordinarios resultados solo se pueden obtener gracias a la implementación de filtros oportunos, como veremos en profundidad en el próximo capítulo.

Por ahora es importante saber que el mejor éxito está asegurado por el correcto equilibrio entre potencia de cálculo y calidad del algoritmo de búsqueda. Buscar una información entre los terabytes (1 TB = 1.000 gigabytes) o petabyte (1 PT = 1000 TB = 1 millón de gigabytes) necesita del empleo de extraordinarios soportes de almacenamiento y formidables sistemas de indexación, con la tarea de identificar tanto en qué punto del enorme archivo se encuentra la información como calcular el tiempo necesario para extraerla.

La Red rebosa de leyendas no siempre verificadas ni verificables a propósito de la capacidad computacional de Google, también porque la empresa revela pocos detalles de su infraestructura tecnológica. Algunas fuentes hablan de cientos de miles de ordenadores conectados entre sí en miles de gigantescos *clusters* que implementan las apropiadas distribuciones GNU/Linux; otras. de superordenadores, dispositivos cuya estética recuerda escenarios de ciencia-ficción: silos enormes superrefrigerados en los que uno o más brazos mecánicos mueven a la máxima velocidad miles de discos duros. Ambas soluciones son plausibles, además de otras, sin estar

necesariamente en contradicción. Lo cierto es que la extrema escalabilidad de las máquinas de Google permite prestaciones excepcionales, puesto que el sistema está «abierto» a continuas mejoras.

De la *brand-identity* a la interfaz participativa

La búsqueda, almacenamiento y recuperación de datos son procedimientos extremadamente complejos y que necesitan, para ser comprendidos, de un conocimiento en profundidad que sobrepasa los objetivos e intenciones de este texto. Más adelante veremos algunos detalles de su funcionamiento. Una atención especial hay que dedicarla a la interfaz, pues mientras que las actuaciones del algoritmo y la arquitectura de la base de datos son elementos estructurales del motor de búsqueda que permanecen visibles al usuario, la interfaz está proyectada y gestionada como imagen del propio Google.

Por interfaz entendemos ante todo el *blank box*[8], ese espacio en blanco en el que se introducen las peticiones o «intenciones de búsqueda» en el marco de la página universal de Google, estudiada para resultar acogedora, cómoda y familiar.

Se trata de un perfil denominado universal, pues está desarrollado en distintos idiomas (actualmente, más de 104, entre idiomas y dialectos personalizables para más de 113 países) y, en cada una de éstos, presenta un modelo de interacción invariable que unifica los comportamientos de búsqueda en un esquema único y homogéneo.

8 En la jerga informática, *black box* hace referencia a una «caja negra» que recibe inputs, los elabora de manera no transparente para el usuario y restituye un output. El concepto de «blank box» reproduce esta forma, pero de manera implícita y por esto ambigua, pues, aún tratándose de un «espacio limpio, vacío, en blanco» (*blank*) está cargado de significados y funciones de búsqueda altamente diversificados.

En la página de Google nos encontramos delante de una interfaz lineal compuesta por elementos esenciales, cada uno con una función bien determinada y universalmente reconocida. Está en condición de aceptar indicaciones de búsqueda de distinta naturaleza y complejidad, desde la introducción de sencillas palabras clave (p. ej., «ippolita») a palabras compuestas, que hay que poner entre comillas (p. ej., «comunidad creativa»), hasta a búsquedas específicas: por ejemplo, las búsquedas pueden ser limitadas a un sitio particular o a un idioma concreto, a páginas procedentes solo de un determinado dominio o a documentos de un cierto formato, y aún más, según el nivel de precisión que se desea obtener. O sea, se trata de un logrado ejemplo de interfaz que alcanza el no fácil objetivo de asociar un significado positivo al espacio en blanco de la página. La interfaz se presenta sin inútiles decoraciones, casi vacía, o mejor rellenada por un único elemento «vacío»: el *blank box*, que da tranquilidad al usuario y tiende a inducir comportamientos activos, en lugar de provocar despiste debido a la ausencia de puntos de referencia o, al revés, por la presencia de inputs visuales excesivos. Se evita así la confusión generada por paginas demasiado llenas, que parecen padecer una especie de *horror vacui*, una ansiedad comunicativa que, en el intento de atraer al usuario con miles de banners, efectos gráficos y juegos, muchas veces consigue el efecto contrario.

No existe una navegación verdadera en la página de Google: las distintas partes de la página tienen un significado funcional, sirven para acceder a los servicios, no para conducir al usuario en un recorrido; su utilización pone en marcha comportamientos que muy rápidamente devienen parte de una rutina de búsqueda, tanto que parecen instintivos al poco tiempo. La interfaz del motor de búsqueda está estudiada de manera que la utilización, la dinámica de funcionamiento y las expectativas del usuario, un usuario genérico, se repitan; de hecho, también después de haber almacenado y digerido las «personalizaciones» del propio usuario, las prácticas de

búsqueda permanecen esencialmente idénticas, hasta el punto de que podemos hablar de un instrumento «universal».

La disposición de textos e imágenes es lineal y aprovecha elementos gráficos recurrentes, por ejemplo, el empleo de colores elementales; las imágenes utilizadas son cualitativamente homogéneas. El estilo de construcción de la interfaz es sobrio, casi desnudo y, en contraste con el diseño de moda de las *brand-identities* (y de las *corporate-identities*)[9], orientado a la búsqueda de una especificidad estética, hace referencia a cualidades perceptivas elementales pero muy eficaces en su sencillez.

De esta inmediata identificación visual procede una facilidad de uso netamente superior respecto a los motores de búsqueda de la competencia. El nivel de ergonomía alcanzado es sorprendente: Google no necesita presentarse como un concentrado de servicios a través de su interfaz; en otras palabras, su arquitectura visual es aquella típica de los portales multiservicio. Las interfaces de los distintos servicios son autónomas y sustancialmente independientes, todas caracterizadas por la presencia de la *blank box* y no linkadas entre ellas de forma directa. Por ejemplo, se necesitan muchos pasos no intuitivos para alcanzar el servicio de code.google.com, pensado para técnicos de variado nivel, partiendo del servicio base de búsqueda de imágenes (images.google.com), dirigido a un público más genérico: es necesario bucear «en profundidad» en el sitio google.com y saber qué buscar. A pesar de esta fragmentación, todos estamos en condiciones de reconocer la red de servicios ofrecida por Google; además, los usuarios pueden utilizar de manera

9 La ideación de una nueva imagen para un producto o un servicio se conoce como *brand identity*; cuando se refiere a una sociedad específica se habla de *corporate identity*. Hoy en día el concepto de *brand* ha superado ampliamente la idea de «sello de distinción», llegando a configurarse como una «marca» que necesita expansión psíquica, territorial, comercial. Para una primera introducción, véase http://es.wikipedia.org/wiki/Marca.

combinada e integrada los recursos informativos puestos a su disposición, ya se trate de aquellos que se limitan a la simple utilización del navegador, de los Google-adictos o de los *Google-totally-adicted*[10] que se vuelcan entusiasmados en cada nuevo servicio.

Esta desterritorialización de los servicios genera un peculiar mecanismo relacional: los usuarios no conocen las nuevas secciones directamente desde Google, sino desde la red informal de los usuarios, desde otros sitios en los cuales los visitantes exponen sus gustos y discuten sus costumbres. El mismo usuario localiza automáticamente la amplia gama de servicios ofrecidos por Google en el momento en que se interesa por un nuevo servicio. Por otro lado, es fácil enfocar las tipologías de usuarios a los que se dirige un servicio y evaluar el nivel de preparación técnica requerido o el nivel de afinidad con otros usuarios. El mecanismo del boca a boca se vuelve algo parecido a un «PageRank(™) relacional».

En una primera aproximación, existen una dimensión relacional local, en la que el boca a boca se produce entre amigos y conocidos, y una dimensión relacional tipológica, en la cual un cierto tipo de usuarios, identificables en base a parámetros estadísticos (edad, sexo, empleo, etc.), utiliza un determinado servicio y pone en marcha la economía relacional.

Los diez problemas relativos a la usabilidad de los sitios web, discutidos por Jacob Nielsen[11], uno de los más conocidos estudiosos de interfaces-usuario, parecen no afectar lo más mínimo el sitio de Google que, aun estando escrito en lenguaje HTML totalmente fuera de estándar[12], consigue asegurar la total visibilidad en todos los navegadores, ya sean gráficos o textuales.

La limpieza gráfica de las páginas es resaltada por una optima gestión visual de los aspectos comerciales. Ningún link publicitario en la homepage o en las páginas de documentación e información: la publicidad en Google se encuentra solo entre los resultados de las búsquedas, separada adrede de los resultados propuestos, pero no ajena a los argumentos buscados. Por tanto, se puede decir que Google es capaz de expresar, al menos en la disposición escénica de sus interfaces, el correcto compromiso entre respeto a los usuarios y necesidad de retorno económico. La publicidad, fuente principal de ingresos de Google, se programa y realiza para que no sea nunca invasiva y para no distraer a los usuarios de su utilización de los servicios.

Los links publicitarios están patrocinados de forma dinámica para seguir el recorrido de un usuario dentro del motor de búsqueda y, por tanto, en segunda instancia, en los sitios de Internet.

Los links comerciales no son estáticos, se modifican y acompañan las búsquedas de los usuarios; esto es posible también gracias a los *feed RSS* (acrónimo de *RDF Site Summary*, o de *Really Simple Syndication*), uno de los formatos más utilizados para la distribución de contenidos web y, en virtud de las diferentes fuentes informativas digitales (diarios, revistas, agencias de prensa, etc.), capaz de modificar dinámicamente la homepage de Google. De hecho Google

10 La Google-manía se extiende y genera nuevos lenguajes; para una panorámica véase: http://en.wikipedia.org/wiki/Google_(search engine); una lista de servicios y herramientas relacionados con Google: http://en.wikipedia.org/wiki/List_of_Google_services_and_tools.

11 Jacob Nielsen, cincuentón informático danés, es una de las mayores autoridades en el campo de la usabilidad de la Web. Nielsen es conocido, entre otras cosas, por sus críticas al exceso de gráficas y animaciones (por ejemplo, *flashs*) que padecen muchos sitios populares y que van en detrimento de la usabilidad, práctica dañina sobre todo para los discapacitados. Su último libro publicado es Jacob Nielsen y Marie Thaie, *Homepage Usability*, Apogeo, Milán, 2002. Los diez problemas de la usabilidad: http://www.shinynews.it/usability/1005-errori.shtml.

12 HTML (acrónimo para *Hyper Text Mark-up Language*) es un lenguaje utilizado para describir los documentos hipertextuales disponibles en la Web. No es un lenguaje de programación, sino un lenguaje de *mark-up*, o sea, que describe el contenido textual y no textual de una página web. Véase: http://es.wikipedia.org/wiki/HTML, y sobre todo el sitio de la W3 acerca del estándar HTML: http://www.w3.org/MarkUp/.

pone su homepage a disposición de los usuarios registrados, haciéndola totalmente configurable gracias a la incorporación de feed RSS: es así posible configurar las previsiones del tiempo automáticas para las ciudades que se desea monitorear, o recorrer el archivo histórico de las búsquedas realizadas. Se pueden organizar los marcadores y los últimos mensajes de correo recibidos, y también tener controlados los archivos del propio ordenador sin solución de continuidad respecto a los contenidos web, gracias a la aplicación de Google Desktop.

El mecanismo de promoción publicitaria, los servicios y los sofisticados mecanismos de perfilación del usuario parecen constituir un todo a nivel estético y de contenidos; por otra parte, los links patrocinados no serían en este sentido nada más que simples sugerencias, gráficamente compatibles y conceptualmente alineados con la operación de búsqueda que se está llevando a cabo. La economía de Google está altamente integrada con la interfaz, hasta el punto de poder ser excluida a nivel visual por quien no esté interesado y, por el contrario, ser aprovechada por quien encuentra interesantes los links y recorridos comerciales propuestos.

También Yahoo![13] y muchos otros motores de búsqueda y portales ponen a disposición herramientas parecidas para la personalización de la propia homepage; pero la cantidad y calidad de las ofertas de Google, de momento, es insuperable. Se trata de configuraciones bastante sencillas que, sin embargo, requieren una cierta familiaridad con las interfaces web y un poco de tiempo para ser realizadas. En el ámbito de la Web el umbral de atención es notoriamente muy bajo, las páginas se visualizan y abandonan en tiempos muy rápidos, del orden de pocos segundos; por tanto, un usuario que invierte muchos minutos, o decenas de minutos, lleva a cabo unas elecciones que revelan mucho de él y de sus costumbres como consumidor. Estas informaciones, cuidadosamente al-

13 Véase por ejemplo MyYahoo!: my.yahoo.com.

macenadas por la compañía de turno (Google o Yahoo!, quien sea), constituyen la verdadera riqueza producida por el propio usuario, y son fundamentales para proponer bienes y servicios específicos por parte de las empresas anunciantes.

La personalización de las páginas hace que un sitio sea más acogedor; el propio sitio se configura como una herramienta personal en la que el usuario invierte tiempo, eligiendo colores, aspecto, contenidos. Un visitante habitual capaz de configurar su página de inicio es cooptado y se le hace partícipe de la construcción de la interfaz web. Concederle al usuario el poder y el control sobre algunas páginas significa promocionarlo desde simple diana de campañas publicitarias a consumidor «inteligente», y desde luego ésta es la mejor y más sutil manera de crear fidelidad, promoviendo la interacción. Se perfilan ambientes dotados de interfaces participativas y no exclusivas para recibir publicidad cada vez más personalizada a fin de entrar todos juntos en el dorado mundo de Google.

PageRank(™) o la autoridad absoluta de un mundo cerrado

El algoritmo que permite a Google asignar un valor a las páginas indexadas por el spider se conoce como PageRank(™).

Ya sabemos que el funcionamiento de PageRank(™) está basado en la popularidad de una página web, calculada sobre la base del número de sitios que hospedan al menos un link que enlaza con ella. A paridad de número de links, dos páginas web tendrán PageRank(™) diferentes según la importancia de quien las haya linkado: con este mecanismo se evalúa el aspecto «cualitativo» de los sitios.

Tomemos un ejemplo concreto: frecuentemente, al controlar las estadísticas de acceso referidas a un sitio, se encuentran gran cantidad de contactos procedentes de sitios pornográficos. Esto sucede porque Google atribuye un ran-

king en función de los accesos, que a su vez vienen visualizados en las páginas de estadísticas públicas. Existen, por tanto, programas que aprovechan el carácter invasivo de esta lógica de conexión y evaluación de los nodos de una red para subir el propio *rank*; como ocurre con frecuencia, los primeros experimentadores son los sitios pornográficos (así ha sido en el caso de las galerías de imágenes o del comercio online).

En la práctica se utilizan algunos programas que se ocupan de buscar sitios con estadísticas de acceso públicas; entonces se produce un elevado número de peticiones, simulando visitas procedentes de un link falso presente en otro sitio, que en la mayoría de los caso es un sitio pornográfico. Este mecanismo de bombardeo multiplica el número de accesos al sitio en cuestión y, en consecuencia, las estadísticas muestran incrementos que hacen subir sensiblemente el *Google-ranking* del sitio y, en última instancia, también el del sitio pornográfico del que han llegado los links: en fin, una ganancia para todos, al menos en cuanto a la visibilidad.

Este tipo de operaciones no es ilegal (nadie puede vetar hacer peticiones a un sitio Internet) y, gracias a esta práctica, los sitios de estadística pública obtienen un ranking más elevado. Además, dicho mecanismo nos permite ilustrar como también la magia tecnológica del ranking de Google, considerado objetivo y veraz, se ve asociada a los «barrios bajos» de la Red, de poco prestigio, y a prácticas de linking más bien equívocas.

Otras prácticas no ilegales que aprovechan la indexación de Google se conocen como SEO (*Search Engine Optimization*); se trata de un conjunto de actividades para mejorar el posicionamiento de un sitio web en los resultados de búsqueda. La oferta de una posición de primer orden se produce a menudo a través de e-mails de spams procedentes de direcciones poco creíbles, evidentemente traducidas con programas automáticos, que prometen resultados asombrosos: «Nosotros registramos su sitio Internet en 910 motores de búsqueda, registro y catálogo web. ¡Nosotros llevamos su sitio Internet a las primeras posiciones de Google y Yahoo! ¡Inténtenlo! No hay ningún riesgo. Por tan solo 249 € (coste único, sin suscripción) en lugar de 349 €». Por supuesto, Google sigue reivindicando su transparencia: «nadie puede garantizar que vuestro sitio aparezca en el primer puesto en los resultados de Google»[14].

Desde el punto de vista matemático, una consecuencia del PageRank(™) basado en el análisis de los links es la integridad de la base de datos; o sea, la determinación de un espacio circunscrito, aunque amplio, en el cual realizar las búsquedas; en suma, en el mundo de Google existe siempre un recorrido que lleva de una página a cualquier otra presente en la base de datos, es decir, en las redes indexadas.

Por tanto, las búsquedas serán tendencialmente funcionales, evitando al máximo la posibilidad de links rotos (*broken links*) o de informaciones distintas de las anteriormente archivadas, presentes en la memoria oculta (*cache memory*). El problema es que de esta forma se induce a los usuarios a creer erróneamente que Internet es un mundo cerrado, conexo, completo, sin carreteras poco iluminadas o recorridos preferenciales, porque parece que, dada una interrogación, siempre se llega al resultado «correcto».

Esto se debe al hecho de que la visión googoliana de Internet se genera por entero a partir de los recorridos que el spider cumple rebotando de un link a otro. Si una página no se cita en ningún otro sitio, entonces no aparecerá nunca en ninguna interrogación realizada por un usuario, pues el spider nunca tuvo la posibilidad de encontrarla, pesarla y evaluarla. Pero esto no significa de ninguna de las maneras que no existan islas de datos; al contrario.

14 Texto de un e-mail recibido en info@ippolita.net en mayo de 2005. La posición de Google respecto a los SEO puede consultarse en este enlace: http://www.google.com/support/webmasters/bin/answer.py?hl=es&answer=35291. Para profundizaciones técnicas, aconsejamos la guía estratégica del posicionamiento en Google, http://www.google-rank.com/.

Un ejemplo son los sitios dinámicos, en los cuales las funcionalidades ofrecidas están basadas en su totalidad en las elecciones del usuario. Uno de estos sitios es trenitalia.com: recopilando la ficha (*form*) el sitio es capaz de ofrecer en tiempo real los horarios de los trenes, las coincidencias, los recorridos más veloces para llegar a un destino. Google no es capaz de entender las peticiones de este form y por eso no indexa los horarios y los recorridos generados dinámicamente por trenitalia.com. Solo la intervención humana puede superar este escollo. La única solución propuesta por Google es la de incorporar en su interfaz herramientas de redireccionamiento a los sitios de compañías aéreas o ferroviarias en el momento en que el usuario busca un recorrido, destino o llegada.

La integridad referencial propuesta por la base de datos de Google debe ser revisada porque supone un mundo único para todos, cerrado y limitado. Al contrario, trazar un recorrido en una red compleja quiere decir llevar a cabo una exploración que siempre determina resultados relativos y parciales.

El sueño de Google contenedor de todo Internet es una idea demagógica particularmente cómoda, útil para sostener la exhaustividad y fiabilidad de las informaciones disponibles, o sea, todas las características que hacen de Google un «servicio único», un dispensador de verdad. En las búsquedas de la vida cotidiana dicha clausura absoluta es muy útil, pues conduce rápidamente a un resultado; en realidad, nos hace ilusión pensar que la libertad consiste en el hallazgo de una calidad total. Sabemos, por contra, que en un sistema reticular complejo no existen verdades absolutas, sino solo autoridades distribuidas según el recorrido que se quiera hacer o, simplemente, en función del tiempo que estemos dispuestos a invertir en la búsqueda. La calidad depende totalmente de nuestra percepción subjetiva de la aceptabilidad del resultado. Las redes que podemos analizar, apreciar y vivir, son objetos complejos cuyos nodos y enlaces están en constante mutación. Como la tarea de aceptar un mapa de navegación referido a una búsqueda depende en definitiva del usuario,

resulta esencial ejercer la capacidad crítica, ser conscientes del carácter subjetivo de nuestro punto de vista. Para generar el recorrido que nos interesa hace falta aceptar la hipótesis de la existencia de una red cerrada y limitada, un mundo cerrado con únicamente nuestras exigencias personales, reconociendo que solo se trata de una localización subjetiva, ni absoluta ni constante en el tiempo. Explorar una red implica la capacidad de dividir las redes en subredes de análisis, y se corresponde con la creación de pequeños mundos localizados y temporales[15].

En la práctica, en la navegación cotidiana los enlaces casuales son de una importancia primordial: la creación de enlaces nuevos e inesperados no puede ser prevista de ninguna manera por el análisis de los elementos de la Red sugerido por el ranking de Google. Estos enlaces funcionan como «puertas dimensionales» y permiten la disminución y hasta la anulación de las distancias entre dos nodos de la Red.

PageRank(™) o la moneda de la ciencia

Además, el algoritmo de PageRank(™), a diferencia de lo que se piensa comúnmente, no es un invento original de Google, sino que procede de los descubrimientos matemático-estadísticos de Andrej Andreevic Harkov, quien en los primeros años del siglo XX analizó fenómenos estadísticos en los sistemas cerrados, es decir, esos sistemas en los que cada elemento es causa o efecto de otros elementos del propio sistema[16].

Sergey Brin y Larry Page seguramente partieron de esta base teórica, pero las mejoras conseguidas no se han divulgado, excepto la patente depositada por Stanford.

---

15 Para profundizar en el concepto de mundos cerrados localizados (LCW, *Localized Closed World*), véase este panfleto realizado por Andrea Marchesini: http://www.autistici.org/bakunin/doc/reti/index.xml.

16 Para una introducción acerca de las cadenas de Harkov, véase: http://en.wikipedia.org/wiki/Markov_chain.

El mejor ejemplo para aclarar la morfología de este algoritmo es el boca a boca entre amigos y conocidos. En las relaciones interpersonales cuanto más se habla de un determinado evento, tanto más éste adquiere importancia y al mismo tiempo deviene parte de un patrimonio común. Si se limita la difusión de ese evento a un entorno reducido su popularidad será menor. Lo mismo vale para los trabajadores del espectáculo: cuanto más consiguen que se hable de ellos, mayor será su ranking, más se conocerán y más famosos podrán ser (por esta razón existen emisiones autorreferenciales como «La Isla de los Famosos»...). Esta misma lógica Google la aplica a los datos online.

Google anuncia este método de forma muy convincente, propagando la imagen de Internet como una gran democracia, pues el algoritmo actúa como si los links fuesen votos referidos a los sitios: poco importa si se linka para hablar mal o bien, lo que importa es que se hable. La falsa verdad de esta «democracia global» procedente de un algoritmo es evidente para cualquiera: como si la democracia dependiera de la tecnología y no de las prácticas de los individuos.

El origen cultural de esta práctica, como ya hemos comentado[17], deriva del sistema, extremadamente elitista, de la revisión por pares (*peer-review*) por parte de los *referees* de las publicaciones científicas: según este mecanismo, la contribución individual de cada investigador se inserta en una red de relaciones, verificaciones y evaluaciones que permite la transmisión y el control de los resultados de la búsqueda. La democracia global de Google se configura como una aplicación del método científico de las publicaciones a la Red, gracias al algoritmo de PageRank(™), una especie de «*referee* tecnológico» para evaluar de forma objetiva las informaciones de la Web, considerando las preferencias expresadas por el «pueblo de los navegadores» a través de los links, y proponerlas en el orden correcto.

17 Véase el capítulo III.

El paralelismo es muy estrecho: por una parte, las publicaciones científicas ganan peso y autoridad conforme a su posicionamiento en el marco de su campo específico de investigación; dicho posicionamiento se obtiene a través de las citas, o sea, las referencias en la literatura científica. De este modo la investigación científica se garantiza una continuidad, pues cada artículo no nace en el vacío, sino que se pone como el «presente» del largo recorrido de la tradición científica. De forma parecida, el spider de Google interpreta los links de las páginas web como «citas» que aumentan el prestigio, es decir, el ranking de esas páginas.

El elitismo científico, base del sentimiento de respeto temeroso que transmite la «ciencia», curiosamente se fundamenta en la práctica de la publicación: hacer «público» no implica «accesibilidad» o comprensibilidad»[18]. De hecho, «los descubrimientos de los científicos, ya sean teóricos o experimentales, ni son ni pueden considerarse conocimiento científico hasta que se registren de forma permanente», como sostenía en los años setenta el sociólogo Robert Merton[19]. Quizás la afirmación sea demasiado perentoria (la ciencia antigua no se transmitía precisamente de modo público: piénsese en la escuela pitagórica en Grecia, en la distinción entre escritos esotéricos y exotéricos, etc.), pero evidencia correctamente el carácter eminentemente público del conocimiento científico moderno. La comunicación no es por lo

18 La impresión de que la ciencia es demasiado difícil de comprender para cualquiera que no sea un especialista se basa socialmente en todos aquellos que, a partir de su vida cotidiana, se sienten ajenos a ella. Los muros del trabajo técnico parecen infranqueables. La expresión de uso común inglesa: «no es ciencia de cohetes» (*it's not rocket science*), normalmente sarcástica observación hecha a alguien que padece insólitos problemas en el cumplimiento de tareas fáciles, solo es un ejemplo de la manifestación de reverencia pública hacia la intensidad intelectual de la ciencia y de su separación de las actividades comunes de todos los días. Véase a este propósito la actividad del CAE, Critical Art Ensemble, www.critical-art.net.

19 Robert K. Merton, *Ciencia, tecnología y sociedad en la Inglaterra del siglo XVII*, Alianza Editorial, 1984.

tanto un subproducto de la investigación, sino parte integrante de una forma del saber cuyos presupuestos son el carácter acumulativo y cooperativo. La ciencia, al menos desde el siglo XVI, por una parte, está orientada a la consecución de resultados novedosos, que puedan representar un aumento del patrimonio del conocimiento; y, por otra parte, asume como punto de partida los frutos de las investigaciones precedentes. Podemos esbozar una historia de la comunicación científica que evoluciona junto con los media destinados a servirle de apoyo: desde la intensa correspondencia epistolar entre científicos hasta la impresión periódica en revistas eruditas o también la comunicación digital. No es casual que los primeros nodos de Internet fueran centros de investigación académica, que necesitaban comunicar y compartir sus informaciones.

De todas formas, la mutación del soporte no ha producido un cambio sustancial en el método de conexión típico de esta forma comunicativa, que sigue siendo el de las citas. Descritas como «moneda de la ciencia», las citas serían una especie de tributo de los científicos a sus maestros e inspiradores. Más concretamente, ponen en contacto la investigación presentada con las ya desarrolladas por el mismo autor o por otros. Es razonable asumir que el número de citas utilizadas de un determinado trabajo pueda representar una indicación de su importancia o, al menos, de su impacto en la comunidad científica. A lo largo de los años este sistema se ha vuelto materia de estudio específica: el análisis bibliométrico es una disciplina que se vale de técnicas matemáticas y estadísticas para analizar los modelos de distribución de la información y, en particular, de las publicaciones. Actualmente, la bibliometría, y ante todo su más conocido indicador, el *impact factor*[20], se utiliza comúnmente como criterio «objetivo» para evaluar la calidad del trabajo científico llevado a cabo por un

20 Eugene Garfield, «The Impact Factor», en *Current Contents*, n.º 37(25), 1994, pp. 3-8; http://www.isinet.com/isi/hot/essays/journalcitationreports/7.html.

investigador o una institución. Un gran archivo para el análisis bibliométrico ha sido colgado online en 1993 justo en Stanford, la cuna de Google. El proyecto SPIRES (Stanford Public Información Retrieval System)[21] nació en 1974 de la colección de notas bibliográficas acerca de artículos de física de las altas energías registrados por la biblioteca universitaria de Stanford. Limitada a un restringido campo de análisis (la física de las altas energías), SPIRES es una base de datos completa y de acceso gratuito, que permite búsquedas complejas también de las citas, un medio que Brin y Page han sabido aprovechar de la mejor manera para poner a punto el algoritmo de PageRank(™). Pero, además del algoritmo, hay otros refinamientos que hacen de Google una verdadera herramienta de mediación global en la Web.

21 Véase el artículo de Marco Fabbrichesi y Barbara Montolli: «L'evoluzione delle abitudini di citazione nella comunicazione scientifica primaria. Il caso della fisica delle alte energie», http://jcom.sissa.it/archive/01/02/L010201/jcom0102%282002%29L01_it.pdf.

5

Además, otras funcionalidades maliciosas

Los filtros en los algoritmos: bases de datos preconfeccionadas y control sobre los usuarios

La teoría de los grafos[1] es la base matemática para la elaboración de todos los algoritmos en las redes, entre los que figura también PageRank(™). Esta rama de la matemática se ocupa, entre otras cosas, de estudiar los métodos para gestionar, crear y recorrer distintas tipologías de redes descriptibles mediante grafos y sus correspondientes medidas. La teoría de los grafos ha tenido un desarrollo especial a partir de mediados del siglo XX con la introducción de los ordenadores electrónicos[2]. Podemos imaginar geométricamente un grafo como un conjunto de puntos en el espacio y de curvas continuas que conectan parejas de puntos sin cruzarse entre sí. El «grafo» (no confundir con gráfico) se define como una figura constituida por puntos, llamados «vértices» o

1 La literatura sobre los grafos es muy extensa; para un primer acercamiento, véase la categoría homónima de Wikipedia: http://es.wikipedia.org/wiki/Grafos.

2 Pese a que la expresión «ordenadores electrónicos» parezca de una precisión gratuita, no lo es. Los ordenadores actuales han sido precedidos por otros tipos de maquinas destinadas a operaciones matemáticas fuera del alcance normal de la capacidad humana. La primera computadora mecánica fue inventada por el frances Blaise Pascal en 1642 y desde entonces aparatos mecánicos, analógicos, electrónicos y digitales se han sucedido y mutado hasta llegar a lo que hoy entendemos por ordenador. [N. del E.]

«nodos», y por líneas que los unen, llamadas «lados», «aristas» o «arcos»[3].

Una red es un tipo de grafo especial, en la que es posible asignar un valor diferente a cada arco (un «peso»), que se podrá utilizar para generar valores de recorrido distintos. Internet es un grafo y podemos definir de la misma forma el conjunto de las páginas web; el sistema de búsqueda de Google se organiza a partir de esta suposición.

Uno de los elementos fundamentales de los algoritmos en las redes es el factor tiempo en relación con el número de nodos examinados. El tiempo para completar un recorrido (o una búsqueda) que une un nodo de la Red a otro depende del número de los elementos de la Red, y varía siempre entre un valor mínimo y un valor máximo, que pueden ser incluso muy diferentes entre sí, según el algoritmo que se utiliza para generar el recorrido. En la red de la Web cada página es un nodo del grafo global y cada link es un arco. Si consideramos la premisa sobre el factor tiempo, es evidente que los resultados propuestos por Google como respuesta a una pregunta cualquiera (técnicamente una *query* en la base da datos) no pueden proceder de una consulta de «todo» Internet.

El spider de Google está permanentemente ocupado en la ardua tarea de copiar Internet en su base de datos; sin embargo, no es creíble que a cada pregunta el motor de búsqueda consulte puntualmente toda la base de datos para encontrar los resultados más congruentes. El factor clave que permite conseguir resultados casi inmediatos depende de una prudente serie de limitaciones en la selección general, es decir, concretamente, de la aplicación de filtros especiales. Planteada una pregunta al motor de búsqueda, un filtro garantiza la velocidad del resultado final mediante una serie de desvíos y de elecciones estudiadas adrede para limitar el intervalo (*range*) de los posibles bloques de datos a examinar.

3 Más formalmente, dado un conjunto V de nodos y un conjunto E de arcos, un grafo G es un conjunto G = (V, E).

De esta manera Google puede facilitar resultados a las preguntas en tiempos excepcionalmente rápidos. Sin embargo, la búsqueda es tan rápida como poco transparente, o sea, no completamente coherente con los datos presentes globalmente en la porción de Red indexada. Dicho con otras palabras, los resultados de la búsqueda se conseguirán velozmente no solo gracias a la potencia de cálculo disponible, sino y sobre todo porque los filtros permiten reducir el pozo de informaciones del que sacar los datos.

La tarea de un filtro consiste en realizar selecciones drásticas en los nodos de la Red excluyéndolos (o valorizándolos) junto con otras eventuales conexiones asociadas. La aplicación de estos dispositivos tiene el objetivo de excluir (o incluir) bloques completos entre los resultados globales.

Todo esto es posible mediante el uso de bases de datos de búsquedas preconfeccionadas, para las que a preguntas estándares se facilitan respuestas estándares, pero también mediante la personalización del perfil del usuario basada en sus búsquedas anteriores, en base a su idioma, a su posición geográfica, etc. De esta manera, un usuario que realiza habitualmente búsquedas en italiano no preguntará a la base de datos global de Google, sino solamente a la parte en italiano, con un evidente ahorro de tiempo.

Para cantidades elevadas de datos es impensable emplear algoritmos transparentes, o sea, que toquen *todos* los nodos de la Red; es preciso introducir manipulaciones, simplificaciones o reducciones de las posibilidades de análisis. Esto por razones técnicas, de cómputo matemático en sentido estricto, y también, obviamente, por razones económicas. Además, sin pecar de malicia injustificada, en un sistema basado en aproximaciones ya dictadas por filtros se puede fácilmente concebir la inserción de filtros posteriores para añadir o mover a posiciones visibles los resultados comerciales, de pago o de pura propaganda.

Sin embargo, hay que observar que desde el punto de vista de Google los filtros añadidos no están relacionados directa-

mente con un interés económico, ya que no sirven para vender un producto, sino que están vinculados al usuario, a sus costumbres y a sus intereses personales. En efecto, Google vende publicidad, no productos (o solo en una mínima parte, por ejemplo los productos hardware como Google Mini y otros sistemas de indexación para empresas): su interés principal es por lo tanto conseguir datos añadidos en base a parámetros sobre los que poder realizar campañas publicitarias con precisión. La personalización de los resultados en base al destinatario es posible sobre la base de informaciones que el mismo Google facilita y recoge de la forma menos invasiva posible. Por ejemplo, los mailboxs, blogs, hard-disks virtuales y otros servicios semejantes representan otros tantos bancos de datos útiles para trazar el perfil de los usuarios mucho más de lo que nos imaginamos.

Los servicios «extras» puestos a disposición por Google, más allá del simple motor de búsqueda, son pues útiles a la empresa para experimentar nuevas vías, pero también y sobre todo porque desempeñan un papel fundamental de «concentradores de informaciones personales» de los usuarios.

Un ejemplo clamoroso han sido los mailboxs de GMail, una especie de hard disk virtual en Google (hasta ahora, más de 2 GB de espacio online, en constante aumento), ofrecidas a través de un sistema de propagación basado en el propio PageRank(™). Dicho brevemente, cada nodo de la red de Google (usuario) dispone de un peso (un cierto número de invitaciones) y puede usarlo para ofrecer el mismo servicio (link) a sus conocidos. Gracias a este método, todo sujeto implicado está controlado desde el punto de vista del uso del servicio; a la vez, pone a disposición de Google informaciones esenciales sobre sus conocidos y amigos.

En un segundo momento, el mecanismo se propaga a la red de las personas invitadas, que pueden realizar nuevas invitaciones; de esa manera se va creando un grafo de las relaciones humanas entre los usuarios de enorme valor estratégico para la personalización de las inserciones publicitarias. Si

consideramos todas las informaciones que se pueden extraer del envío de e-mails (a quién, por qué razón, en qué idioma, en qué formatos, con qué palabras claves, con qué datos adjuntos, etc.), podemos imaginar que en el banco de datos de Google, además de una copia forzosamente parcial pero significativa de Internet, existe también una copia parcial, y también significativa, de las relaciones humanas, emotivas y profesionales de los usuarios del servicio.

En teoría los filtros sirven simplemente para que la interrogación sea más rápida y más ajustada a las exigencias de los individuos; pero además también son absolutamente necesarios desde un punto de vista técnico. Sin embargo, su uso demuestra que, para un sujeto de hecho hegemónico en el ámbito de los servicios de búsqueda, es extremamente sencillo explotar sus peculiaridades a fin de aprovechar los datos a su disposición en un sentido comercial y lesivo de la privacidad.

En suma: actualmente, la base de datos de Google, mediante algunas palabras claves, es capaz de ordenar los resultados de una búsqueda (*query*) de manera diversificada según el tipo de usuario que tiene que atender, es decir, en base a los datos relativos a aquel usuario que obran en su poder. Lejos de ser resultados «objetivos», los resultados de toda búsqueda son pues calibrados; más aún, el uso de los servicios de búsqueda refina las capacidades del sistema para «reconocer» al usuario y ofrecerle resultados adecuados.

El uso de cada servicio está subordinado a la aceptación de los reglamentos y excepciones por parte de los usuarios, y Google en toda ocasión se compromete a no divulgar informaciones reservadas y personales en general; pero es fácil pensar que Google pueda comercializar o explotar para los más variados fines los datos de los usuarios, sin considerar, por el momento, la posibilidad (sería mejor decir: la certeza) de que militares y fuerzas de policía de distinto tipo puedan solicitar y tener acceso a informaciones semejantes por presuntas razones de «seguridad nacional». Añadir filtros de

búsqueda para personalizar los resultados es en concreto el escenario más factible.

Los *cookies* de Google, galletas que dejan su huella

El dibujo del perfil de los usuarios[4] se basa siempre en un sistema de selección y reconocimiento. En la red de Internet distinguimos entre métodos de personalización explicita e implícita. La explicita necesita de un procedimiento ad hoc de registro, que implica el envío a través de un formulario (*form*) de datos personales por parte del usuario. Las informaciones enviadas, archivadas en la base de datos, son analizadas mediante una serie de parámetros útiles para segmentar en grupos homogéneos la totalidad de los usuarios registrados (edad, sexo, profesión, intereses, etc.). La personalización implícita se realiza a través del trazado de usuarios anónimos durante sus visitas a un sitio, mediante la dirección IP o a través de *cookies*, o sea, galletas. Los *cookies* son pequeños archivos de texto utilizados por los sitios web para almacenar algunas informaciones en el ordenador del usuario. Cada vez que el usuario vuelve a estos sitios, el buscador devuelve las informaciones guardadas en la «galleta». El objetivo es automatizar algunas operaciones de autentificación (*login*), recordar eventuales operaciones en curso, pero sobre todo asociar el visitador con todas las informaciones memorizadas en sus anteriores visitas.

La mayoría de los sitios de Internet que ofrecen servicios online usan los *cookies* y Google no constituye en absoluto

4 Se trata de un sector clave de la economía de la búsqueda, en continua expansión, que incluye conceptos como: *Clickstream analysis,* CRM (*Customer Relationship Management*), *data mart*, *data mining, data warehouse*, OLAP (*Online Analytical Processing*). Para una introducción elemental sobre el argumento, véase: http://www.diodati.org/scritti/2002/g_stat/stat05.asp Es interesante anotar que las técnicas de personalización de los usuarios proceden, en última instancia, de las actividades de *criminal profiling* (perfil criminológico).

ninguna excepción[5]. Gracias a la combinación de *cookies* y filtros añadidos sobre los algoritmos es posible mantener un trazado de la navegación de una persona y acumular informaciones sobre su «huella».

Pongamos un ejemplo: el individuo X posee un número de teléfono móvil registrado a su nombre, con el que llama a su familia, a algunos amigos y a los compañeros de trabajo. Después de un cierto tiempo decide cancelar aquel número y utilizar otro no registrado a su nombre, pensando así estar protegiendo su privacidad. Con el nuevo teléfono restablece su círculo de conocidos, contactando con la propia familia, algunos amigos y los compañeros de trabajo. Aquella secuencia de «conexiones sociales» (familia, amigos, compañeros) resulta ser una secuencia única entre todas las llamadas telefónicas del mundo, vinculadas indisolublemente con el protagonista de este ejemplo. No es imposible formalizar una secuencia semejante con un grafo que representa los nodos y los arcos de una red: los valores (los respectivos «pesos» de los arcos que conectan los nodos) podrían ser asignados asociando a la «cercanía» un grado proporcional de «amistad» con el punto de partida del análisis, es decir, con nuestro hipotético individuo X.

La eliminación de los *cookies* es por lo tanto una óptima práctica de defensa de la privacidad, pero un razonamiento análogo al ejemplo anterior puede fácilmente ser aplicado al uso de los motores de búsqueda. Mediante los *cookies*, buscando informaciones sobre algunos intervalos temáticos específicos, es de hecho posible identificar grupos de personas e incluso directamente individuos sobre la base de su huella unívoca dejada en la Red.

La propia huella que marca nuestros movimientos, nuestros contactos sociales (o telefónicos) es única en cuanto a

5 Para más información sobre *cookies* volátiles: http://www.google-watch.org/cgi-bin/cookie.html; http://www.sgroogle.org/. Google Watch explica la gestión de los *cookies* por parte de Google y métodos para defenderse de la personalización. Scroogle se propone anonimizar las bús-

las preferencias, los gustos, las idiosincrasias y las pasiones que nos distinguen de los demás. Nuestras pasiones son, en este caso, los sitios que visitamos y, sobre todo, las búsquedas que realizamos durante nuestra navegación. Esta cantidad de informaciones que ofrecemos a cualquier motor de búsqueda hace que sea posible la reconstrucción de nuestra «huella»[6].

Como toda galleta que se precie, también las de Internet tienen una fecha de caducidad. Los sitios de Internet que asignan a nuestro navegador un *cookie* están obligados a establecer una fecha de caducidad, es decir, una fecha tras la que el navegador puede proceder a la eliminación de los datos contenidos en el propio *cookie.* No es inútil utilizar de la mejor manera los *cookies*; y en este sentido destaca el hecho de que Google haya sabido explotar para su beneficio un artificio técnico conocido por los desarrolladores de todo el mundo, POSIX (el estándar internacional que permite la interoperabilidad de los sistemas operativos Unix y Unixlike, que incluyen también GNU/Linux). En efecto, la fecha de caducidad de Google es el 2038, es decir, casi la misma fecha máxima programable en un *cookie,* y esto comporta que el navegador de nuestros sistemas operativos no eliminará nunca aquel *cookie* y las informaciones guardadas en él[7].

quedas hacia Google, de forma que el *cookie* eterno de Google no pueda realizar un trazado histórico de las búsquedas realizadas por el usuario.

6 Sobre esta temática, véase la ponencia «La privacy y los motores de búsqueda» de Claudio Agosti «vecna» – s0ftpj.org con ocasión del congreso «e-privacy – fra Trusted Computing e Data Retention», que tuvo lugar en Florencia el 19 y el 20 de mayo de 2006, http://e-privacy.firenze.linux.it/, disponible en la dirección http://e-privacy.firenze.linux.it/atti/e-privacy_2006_Agosti_Privacy_e_motori_di_ricerca.pdf.

7 Véase a este propósitio el «Apéndice I: El fin del mundo en una galleta».

Onanismo tecnológico: crea, recrea y consume tus contenidos

Es imposible seguir las rapidísimas evoluciones e innovaciones *made in Google*: nuevos servicios se suceden de forma casi convulsiva y se hace muy difícil entender cuáles son los realmente destinados a tener un impacto en nuestras vidas y cuáles, por el contrario, serán olvidados en pocos meses o semanas. Por otra parte, tampoco es interesante o útil perderse en elaboraciones complejas, en taxonomías exhaustivas que inevitablemente dejan de lado excepciones significativas, si consideramos el elevado grado de caducidad de las informaciones y de las innovaciones en Internet. La naturaleza fluida y dinámica de la Red desanima, si acaso alguien todavía sintiera la necesidad de cualquier enfoque totalizador y omnicomprensivo que se revela veleidoso en el momento mismo de su farragosa exposición.

Sin embargo, es posible esbozar, desde una perspectiva subjetiva y ubicada, una visión del fenómeno Google en toda su complejidad, sin adentrarnos en las técnicas específicas y tampoco en los pronósticos de futuro estériles. Resulta sin duda relevante la figura destacada del *prosumer*, sobre todo en lo referente al concepto de personalización.

Google es conocido por la costumbre consolidada de conceder versiones betas, no definitivas y todavía en fase de test, de sus servicios: esta dinámica, como hemos visto en el capítulo precedente, se sirve directamente del apoyo de las comunidades de desarrollo free software. Los usuarios, dando su *feedback*, sugerencias e impresiones de uso, contribuyen de manera determinante a la consolidación o no de un nuevo servicio: a la vez son productores y usuarios de los servicios que usan, figuras híbridas llamadas por eso *prosumers*[8].

Colocándose como sujeto de mediación global de los contenidos de la Web, Google vende esencialmente tecnologías y resultados de búsqueda (mediante la publicidad) a los usua-

8 *Productor-consumer*: productor-consumidor.

rios que, por un lado, tienden a ser creadores de los contenidos de la Red y, por el otro, los consumidores de aquellos mismos contenidos que han modelado en primera persona a través de los servicios de Google. Dos ejemplos, aparentemente sin vínculo entre ellos, pueden esclarecer este ciclo cerrado de producción y de consumo de los contenidos: el Google Web Toolkit (GWT)[9] y la convergencia entre GTalk y Nokia[10].

Google lanzó en mayo del 2006 el Google Web Toolkit, un framework que permite desarrollar aplicaciones AJAX mediante el lenguaje Java. AJAX (*Asynchronous JavaScript and XML*) es una técnica para desarrollar aplicaciones web interactivas y dinámicas, usando una combinación del clásico HTML (o XHTML) con CSS para la parte visual, y JavaScript para mostrar dinámicamente las informaciones e interactuar con ellas. En resumen sitios extremadamente dinámicos en los que no es necesario esperar cada vez a que se carguen todas las informaciones de la página. Gmail, por ejemplo, utiliza AJAX. La novedad es significativa porque de este modo cambia el enfoque de la creación de las aplicaciones web, que se escriben en un lenguaje de objetos de alto nivel, Java, luego recogidas por GWT y compatibilizadas con todos los navegadores. Se ha pretendido presentar estos avances técnicos como los precursores de la web 2.0, pero la realidad es que las tecnologías para la creación de software multiplataforma, para compartir los *bookmarks*[11], el social networking, la agregación automática de datos, etc., son conocidas desde hace años. En este sentido, es notoria la hipocresía de multinacionales como SUN, que alardean de la entrada en la Era de la Participación vendiendo como una revolución las innovaciones tecnológicas generadas por organismos como el W3C[12] (por ejemplo la estandarización mediante XML/RDF), a la vez que ignoran la

9 http://code.google.com/webtoolkit/.

10 googlesystem.blogspot/2006/05/nokia-770-will-include-gtalk.html.

11 *Bookmarks*: directorios en los que se almacenan las direcciones web utilizadas habitualmente por los usuarios. [N. del E.]

12 Consorcio en el que figuran multinacionales como Adobe, France Telecom, IBM o la propia Google, dedicado a unificar criterios y estándares tecnológicos para Internet. [N. del E.]

disposición a la cooperación que, desde hace decenios, es una de las señas de identidad de la cultura hacker[13].

Desde luego AJAX y tecnologías parecidas solucionan de raíz el problema muy común de la portabilidad de los sitios web, que actualmente son difícilmente visualizables por todos los navegadores. El código del framework está disponible bajo licencia Apache, es decir, fundamentalmente open source, pero, como acontece con las iniciativas de code.google.com, algunas partes fundamentales (en este caso, el compilador Java-to-JavaScript y el *hosted web browser*) se distribuyen solo en binario y es preciso suscribir una licencia ad hoc, que prohíbe redistribuirlos, derivar de ellos ulteriores productos e incluirlos en productos comerciales. Además, cada vez que se usa el *hosted web browser*, que permite comprobar las aplicaciones en la propia máquina antes de publicarlas en Internet, se contacta un servidor de Google, oficialmente para comprobar que se está usando la versión más actualizada del producto. Aunque el código producido se puede distribuir libremente, incluso para fines comerciales, es evidente, sin embargo, que éste es un método de control extremadamente eficaz aplicado a los desarrolladores en lugar de a los usuarios.

GWT es en la práctica un instrumento simple para crear sitios perfectamente compatibles con los sistemas de indexación de Google. Actualmente, es preciso conocer un lenguaje no sencillo como Java, pero no es impensable imaginar el desarrollo de instrumentos que permitan incluso a un usuario novato poner en la propia página web objetos como barras de las herramientas, galerías de imágenes, menús de varios tipos, etc., sin escribir ni una línea de código. Naturalmente ya exis-

13 Sobre la implementación de la web semántica, cosa distinta de la «web 2.0», véase la *roadmap* de 1998 de Tim Berners-Lee, http://www.w3.org/DesignIssues/Semantic.html; sobre la coordinación del W3C, http://www.w3.org/2001/sw/; un documento introductorio a la filosofía de la web semántica: http://www.autistici.org/bakunin/doc/reti/index.

ten programas extremadamente sencillos para crear sitios web (WYSIWYG, *What You See Is What You Get*); pero GWT es vía web. Estos contenidos estarían pues inmediatamente listos para ser usados en dispositivos fijos o móviles de toda clase, siempre y cuando dichos dispositivos tengan acceso a la Web.

Imaginemos ahora que Google firme acuerdos comerciales para la construcción de hardware a medida, ofreciendo a quienquiera que use sus servicios instrumentos sencillos para construir páginas web visibles en los PC, PDA, etc., e indexadas de la manera más fácil posible por sus spiders. Google de hecho no proporciona programas bajo pago de licencias, como Microsoft: necesita, como sabemos, difundir sus propios estándares para poder gestionar más fácilmente su economía de búsqueda.

Y así nos encontramos con el acuerdo entre GTalk y Nokia. GoogleTalk es el servicio VoIP[14] de Google, recién integrado en la cuenta de correo de GMail, de modo que los que forman parte de la «Google community» puedan no solo enviarse correos, sino también chatear y hablar en tiempo real. A partir de finales de mayo 2006 está disponible para los nuevos dispositivos móviles de Nokia, llamados *Internet tablet*, una clase de móviles evolucionados pensados para navegar por la Web. Así Google entra por la puerta grande en el mundo de la telefonía móvil, con la perspectiva de integrarse bien pronto en las redes públicas inalámbricas (*wireless* o *wimax*) que empiezan a ser instaladas en distintas ciudades, aeropuertos, áreas de servicio en las autopistas, etc. Una perspectiva de convergencia con el vídeo también parece plausible: vídeo.google.com es una mina de películas, y la televisión en los móviles es el próximo paso.

En otras palabras, Google, facilita los instrumentos para crear contenidos según sus estándares. Es la personalización extrema, el equivalente al mecanismo de «largas colas» (que

14 Véase la nota 10 del cap. 2. No solo Skype y Google ofrecen servicios VoIP, también lo hacen programas *free* como Asterisk, http://www.asterisk.org/.

permite ofrecer exactamente el producto personalizado a cada consumidor) en el campo de la creación de contenidos web: el usuario crea «exactamente» lo que desea en el formato estándar de Google. La descentralización total a nivel de creación de contenidos va paralela a la descentralización total de la publicidad —ya *ad personam*— y, por lo tanto, del suministro de productos «personalizados».

Un sistema invasivo para imponer un estándar, pero formalmente «democrático» porque está en manos de los usuarios, a tiro de clic, a tiro de navegador. Lo que se vende como democracia electrónica se convierte en una estandarización que permite digerir los contenidos creados por una miríada de usuarios y asociarles la publicidad más adecuada.

Los navegadores como ambientes de desarrollo

La explosión de nuevos y cada vez más potentes servicios web desde 2005 está convirtiendo el navegador de un instrumento de simple navegación a un instrumento de desarrollo. Un amplio abanico de tecnologías altera los estándares de programación web corrientes, ofreciendo a los desarrolladores un instrumento fácil, multiplataforma, completo y fiable: el navegador mismo.

En los últimos años se ha ido moldeando lentamente un nuevo estilo en la creación de los sitios de Internet, también gracias al mayor cuidado de la portabilidad y la accesibilidad de los contenidos: claro ejemplo de ello son la difusión de las hojas de estilo (*Cascading Style Sheet,* CSS estándar y CSS2) en lugar del HTML puro y duro de los validadores, del propio estándar XML. Los diseñadores gráficos y diseñadores de webs encuentran unos aliados óptimos en los navegadores, cada vez más sofisticados y respetuosos de los distintos estándares, que les permiten realizar sitios visualizables en los dispositivos más diversos, manteniendo e incluso ampliando la libertad de expresión en el terreno técnico.

En especial, el nacimiento de navegadores como Mozilla Firefox y su rápida difusión han permitido una fuerte interacción entre los desarrolladores de sitios de Internet y los desarrolladores de este navegador, consiguiendo en poco tiempo sanar casi todos los *bugs* y las incompatibilidades entre los estándares web. La incompatibilidad entre Internet Explorer, Opera y muchos más navegadores, propietarios o no, es un problema bien conocido por cualquiera que haya creado páginas web. La sinergia desarrollada por Mozilla, aparentemente simple y tal vez obvia, constituye una novedad absoluta en la historia de los navegadores. Otra característica interesante de los productos Mozilla es la estructura modular desarrollada alrededor del motor de *rendering* Gecko, que permite añadir cualquier funcionalidad imaginable: entre las más conocidas y difundidas figuran instrumentos para eliminar la publicidad en las páginas web, para tener siempre a la vista las previsiones del tiempo o la situación del mercado bursátil.

Los navegadores son, en resumen, instrumentos cada vez más fiables, que permiten la creación de complejos sitios de Internet y asumen las características de verdaderos y auténticos programas: tienden incluso a sustituir las aplicaciones de uso común. Un ejemplo ilustrativo son los diversos instrumentos para oficina propuestos por Google como alternativa web a la suite de productividad de Microsoft Office y también de OpenOffice[15]. Actualmente, es posible utilizar Writely (producido por una sociedad adquirida) como procesador de textos, Google Spreadsheets como hoja de cálculo, Google Page

15 También existen suites completas para oficina utilizables vía web, como gOffice, producido por Silveroffice, Inc. de San Francisco, que sostiene: «gOffice no es una filial de Google, Inc., pero nos gusta pensar que será una bella lucha puesto que su misión seguirá siendo la de indexación de todas las informaciones del mundo, mientras que la nuestra es la de ayudar a la gente a crear esas informaciones. ¡Necesitamos pues de Google para buscar los contenidos que nuestros usuarios crean con gOffice! Tal vez tendríamos que añadir un botón a nuestra barra de herramientas llamada "búsqueda"... www.goffice.com».

Creator para crear páginas web rápidamente. Todos los servicios están en fase de *betatesting* con invitaciones: es preciso tener una cuenta Google para utilizarlos, evidentemente bajo estricto control.

Los desarrolladores, por su parte, desplazan su interés cada vez más hacia la Web, también gracias al uso de instrumentos como GWT. Naturalmente Microsoft no se queda mirando y tiene ya lista la versión beta (sumándose a la estrategia de «*betatesting* difuso» de Google, procedente como sabemos de las prácticas del free software) de Office System (o sea Office 2007), que prevé una integración cada vez más acusada respecto a las herramientas web, pese a seguir siendo una aplicación a instalar.

Los navegadores se están convirtiendo en ambientes de desarrollo completos para la creación de contenidos estándar, es decir SDK (St*andard Development Kit*). ¿Dónde está realmente la innovación que ha hecho posible transformar los navegadores en SDK? Sin duda, podemos hablar de un nuevo paradigma de la programación: se pueden crear ahora programas realmente multiplataforma, distribuidos, lado client, residentes en los server, fácilmente actualizables y sin necesidad de complejos framework a instalar en los ordenadores de los usuarios. Los contenidos, incluidos los datos personales de los usuarios, son almacenados cada vez más en máquinas remotas (por ejemplo los servers de Google), a los que se accede vía web, es decir, mediante navegadores.

A menudo la elección de navegadores open source, como Mozilla, se debe a su sencillez de configuración, y a la disponibilidad de numerosas extensiones potentes y gratuitas. Los desarrolladores se basan en estos navegadores para implementar programas cada vez más complejos y estructurados. El nacimiento de programas basados solamente en la Web provoca dos fenómenos, ambos con fuertes repercusiones en el mercado y en los usuarios: los programas binarios a instalar devienen obsoletos y los navegadores se convierten en

programas más complejos, a menudo fuertemente modulares, y más apetecibles en el mercado de las tecnologías de la información. Por lo tanto, en el futuro podemos esperarnos menos programas .exe (Windows), .dmg (Macintosh) o paquetes Unix instalables, y más instrumentos integrados en los navegadores, desde las extensiones para leer los feed RSS al podcasting, desde Gmail a softwares completos para oficina.

El control capilar que puede ser realizado por parte de los suministradores de los servicios web hace que esta dinámica sea potencialmente peligrosa para todos nosotros. Quien ofrece los servicios, en efecto, conoce exactamente la identidad digital de los propios usuarios, el tiempo de uso del software y los contenidos en elaboración, ya que controla todo acceso y detalle relativo al uso de los instrumentos que pone a su disposición.

Se trata, a nivel técnico, de una interacción continua entre el client (navegador) y el servidor (el que ofrece los servicios), en la que este último es capaz de monitorear constantemente las solicitudes, los tiempos y los resultados finales.

Muta también el escenario de la seguridad: con la excusa de ofrecer productos «a salvo» de las garras de crackers y hackers de toda clase, los mecanismos de autentificación no se encuentran ya en el programa compilado y concedido sin fuentes, sino que están directamente en los servidores de los proveedores de servicios y productos. Por lo tanto, para «entrar» en un software ya no basta con desensamblar un código en el propio ordenador y comprender como funciona, sino que es preciso «perforar» servidores remotos.

Privacidad, paranoias y poderes

La estrategia de la acumulación ha hecho de Google un sujeto capaz de poner en dificultad hasta a Microsoft, empezando una guerra no exenta de golpes por la estandarización y el control del acceso a la Web y a las redes que usamos cada

día. En cuanto sujeto de mediación global de informaciones, el fenómeno Google tiene que ver directamente con todos los usuarios de informaciones digitales, es decir, todos nosotros. Recorrer la historia de Google significa por esto re-recorrer nuestra historia de exploradores de Internet y de la Web, pues demasiado a menudo hemos delegado la gestión de nuestras informaciones, nuestros sitios, galerías de imágenes, cuentas de correo, blogs, sms, conversaciones telefónicas, etc., a empresas que lo son todo menos desinteresadas.

La estrategia de la objetividad —que Google actúa poniendo el acento en la investigación, en la excelencia académica, en la superioridad tecnológica, en el cuidado estudio de las interfaces— es un velo que oculta la inquietante perspectiva de un único punto de acceso a los datos producidos por los ingenuos usuarios.

La estrategia del open source, en definitiva, Google la necesita precisamente para cooptar el método de desarrollo cooperativo de las comunidades digitales, haciendo que sea funcional a la propia «misión». También en este caso hemos visto como Google se mueve con extrema desenvoltura, proponiendo métodos «nuevos» para explotar dinámicas meritorias: *Summer of Code* es el ejemplo más evidente.

La actividad de Google, en resumen, constituye un peligro evidente para cualquiera que esté interesado en las temáticas de la privacidad y, en un sentido más amplio, en la construcción consciente del propio álter ego digital. Se trata del surgimiento de un conglomerado de poder con una fuerte influencia ya hoy en día sobre la vida de demasiados individuos. Google tiene informaciones reservadas que analiza sin descanso para promover una personalización cada vez más cuidada del cáncer publicitario. Además, ya que la acumulación de poder generalmente favorece la ambición de dominio, sería oportuno indagar en profundidad este fenómeno.

No existen respuestas globales capaces de solucionar una vez por todas el problema de la privacidad. El Gran Hermano no existe o, cuando menos, como toda paranoia que exclu-

ye horizontes de liberación y oscurece posibles vías de fuga, es útil y funcional al ejercicio del poder dominante.

Ocultar, criptografíar y estenografiar son prácticas útiles, pero no son soluciones definitivas: el objetivo y el deseo continúa siendo la comunicación y el compartir, que solo permite la «publicidad», es decir, el movimiento de «hacer público». La obsesión por la privacidad cae muy rápidamente en la paranoia del complot; en este sentido, no es útil entregarse a la construcción de complicadas alternativas para crear redes absolutamente seguras e impenetrables. La tecnología es una ocasión de apertura y participación, ya que usar las máquinas significa usar criaturas híbridas, artefactos materiales (pertenecen en ese sentido al mundo natural) investidos de valores y significados culturales (pertenecen al universo de la «cultura»). Las redes son expresión de una dinámica coevolutiva de máquinas mecánicas, biológicas y significantes: la tecnología es mestiza por nacimiento. Crear redes significa de hecho conectar máquinas de varios tipos; significa crear métodos para compartir, métodos de traducción, métodos de intercambio; no es posible quedarse encerrados en sí mismos, es preciso meterse en el juego y cambiar.

Son necesarios investigación y análisis competentes; la denuncia de los mecanismos del dominio tecnocrático hace más falta que nunca. Renunciar al pensamiento crítico equivale a sucumbir a las paranoias del control, que sí se hace cada vez más difuso. Es posible utilizar en sentido paradigmático la historia de Google para trazar líneas de fuga no ingenuas, para imaginar prácticas de autogestión de las tecnologías. Google es de hecho el punto de encuentro entre el estilo meritocrático de la Academia, el empuje a la innovación a toda costa y el capitalismo financiero más avanzado.

Ha llegado el momento de desarrollar redes autónomas y descentralizadas para contraponer el deseo de «explorar» y «recorrer» las redes a la necesidad de «acceder» a los datos, para centrar la atención en el camino en lugar de en el resultado.

6

Calidad cantidad y relación

La emergencia de la información

La sociedad de la información presenta una extrema heterogeneidad de manifestaciones: utiliza al mismo tiempo y además con total desenvoltura sistemas de comunicación reticulares (telefonía), versiones digitales de *media broadcast*[1] tradicionales nacidos fuera de la Red (diarios, radio, TV. online, etc.), instrumentos de comunicación universales nacidos con la Red (e-mail) y sistemas distribuidos (P2P). En cualquier caso todos estos sistemas se basan en una única materia prima: la información. Dentro del especifico ámbito de los motores de búsqueda, o sea, del descubrimiento de informaciones (*information retrieval*), podemos afirmar que la información es el conjunto de las páginas que se publican en la Web[2].

---

1 Un sistema de transmisión *broadcast* envía información a muchos sistemas receptores de forma unidireccional (comunicación de uno a muchos). En una red de ordenadores, una transmisión *broadcast* envía informaciones a todos los ordenadores en conexión a esa red; otros sistemas de transmisión son *unicast* (un solo receptor), *multicast* (muchos receptores), *anycast* (cualquier receptor dentro de un grupo). Los sistemas P2P, precisamente como los sistemas telefónicos, son justo lo contrario de los sistemas *broadcast* porque permiten una comunicación del tipo muchos-muchos.

2 No es este el lugar para un exhaustivo análisis del concepto de «infor-

El crecimiento cuantitativo y cualitativo de estas páginas y de sus contenidos ha sido y sigue siendo desmedido. Esto depende de la facilidad con la que es posible generar nuevos contenidos para la Red. Los contenidos no son islas de datos sin conexión, sino que se constituyen en la variedad de las relaciones y de los links que enlazan las páginas, los sitios, los documentos, los mismos contenidos entre sí.

El acceso directo, sin mediación, a esta masa de informaciones es absolutamente imposible, hasta como mera hipótesis: sería como sostener que se puede hojear la Web «a mano». Por esto existen los motores de búsqueda, herramientas estudiadas para filtrar la complejidad de la Red y actuar como interfaces entre nosotros y la información, satisfaciendo las búsquedas.

En los capítulos precedentes hemos repasado los instrumentos fundamentales para el funcionamiento de un motor de búsqueda, o sea, los medios de los que se sirve Google y los otros para devorar las páginas web, analizarlas y catalogarlas en base a los algoritmos de ranking, para archivarlos después en dispositivos hardware especiales y mostrarlos al usuario de acuerdo con sus búsquedas.

La cantidad de páginas memorizadas adquiere entonces un valor crucial para evaluar la potencia técnica y económica de un motor de búsqueda. Cuanto mayor es la amplitud del patrimonio de las páginas examinadas, tanto mayor será la potencial profundidad y fiabilidad de las búsquedas efectuadas, por supuesto, en el límite de un contexto específico.

Pero, aunque la extensión del patrimonio de un motor de búsqueda pueda ser enorme, nunca podrá ser completa y total, con independencia de cuánto tiempo, dinero y tecnologías se inviertan en él. Es absurdo pensar en poder conocer, o más banalmente copiar y catalogar, todo Internet: sería como pretender conocer la totalidad de un mundo vivo y en constante mutación.

mación», aunque esta definición está simplificada adrede. Por ejemplo, no tiene en consideración que los motores de búsqueda como Google indexan también otros contenidos que no son de la Web, como el archivo de los mensajes Usenet o el contenido de los ordenadores que utilizan Google Desktop; de todas formas, puesto que esta masa de informaciones se presenta a través de una interfaz web de mediación, consideramos aceptable la definición adoptada.

Los dispositivos de almacenamiento de las informaciones de los motores de búsqueda como Google se parecen a unos vasos. Imaginemos tener que llenar un vaso de gran capacidad con esas diminutas gotas que son las páginas web que constituyen las informaciones de la Red. Si consideramos que el vaso es capaz de contener la totalidad de las gotas (la totalidad de la Red), nuestra tarea consistirá en intentar capturarlas todas de manera sistemática y repetitiva.

Si, por el contrario, pensamos que las gotas son demasiadas para nuestro vaso y que su recogida no será ni previsible, según los algoritmos, ni rápida, sino que potencialmente podría resultar una tarea ilimitada, deberemos cambiar de táctica: también porque las informaciones varían con el tiempo, las páginas se modifican, los recursos cambian de dirección...

Acaso podríamos decidir elegir tan solo las gotas más grandes, o concentrar nuestros esfuerzos de recogida en puntos donde caen en mayor cantidad o, más simplemente, podríamos recoger aquellas que nos interesan más y luego intentar conectarlas entre sí de la forma que consideremos más interesante.

Aun cuando los motores de búsqueda continúen poniéndose el objetivo de catalogarlo *todo*, es más correcto aplicar a la Web una praxis localizada, o sea, asumir que para cada «intención de búsqueda» existen muchas respuestas posibles, entre las cuales algunas son «mejores» puesto que cumplen con determinados requisititos de rapidez y exhaustividad. Debemos recordar siempre que la calidad de las respuestas depende de nuestra subjetiva percepción de la aceptabilidad del resultado. Para poder aceptar o rechazar un elaborado de búsqueda resulta esencial el ejercicio de la capacidad crítica, la conciencia de la subjetividad del propio punto de vista.

Para generar el recorrido que de verdad nos interesa analizar es necesario suponer la existencia de una red finita y limitada, de un mundo cerrado por nuestras exigencias personales, sabiendo de todos modos que se trata de una localización subjetiva, ni absoluta ni constante en el tiempo. Desde el punto de vista analítico, explorar una red implica la capacidad de dividir las redes en subredes de análisis, y se corresponde con la creación de pequeños mundos localizados y temporales (LCW, *Localizad Closed World*), en los que existe al menos una respuesta a las búsquedas efectuadas. En caso contrario, muchas búsquedas no finalizarían nunca, entre otras cosas porque los datos que pueden ser analizados son muy superiores a las capacidades de elaboración de un sujeto humano. Dicha respuesta podría no ser satisfactoria, pero cambiar o especificar la búsqueda, enriquecer el propio punto de vista entonces podrá generar un recorrido más coherente con nuestras exigencias. Considerar la Web un mundo cerrado localizado significa también reconocer que la dinámica misma con la que las informaciones nacen, germinan y se distribuyen en la Red (aunque no se vuelvan a actualizar) es un fenómeno «emergente», ni causal ni tampoco casual. La emergencia[3] es un comportamiento descriptible en términos matemáticos, una explosión de complejidad inesperada e imprevisible, pero sobre todo es un evento en condiciones de determinar situaciones no descriptibles de manera exhaustiva. Analizar y recorrer un «universo emergente» como la Web implica por tanto un continuo reposicionamiento personal que determina un «mundo cerrado localizado»

3 La emergencia es un principio que describe el comportamiento de los sistemas complejos. Se puede definir como el proceso de formación y emersión de comportamientos y esquemas complejos a partir de reglas más sencillas. Las ciencias del conocimiento utilizan el concepto de emergencia para describir «la explosión» de la complejidad típica de la inteligencia: partiendo de elementos sencillos inconscientes de por sí (las neuronas), se genera el pensamiento, la conciencia manifiesta de los procesos cerebrales: http://es.wikipedia.org/wiki/Emergencia_%28filosof%C3%ADa%29.

de competencias y expectativas, y también la abertura hacia nuevas posibilidades de exploración (hay otros mundos siempre posibles fuera del propio mundo cerrado) y, por tanto, la conciencia de poder alcanzar resultados siempre inevitablemente parciales.

Cantidad y calidad

La indexación por acumulación de páginas es un fenómeno cuantitativo que en sí mismo no determina la calidad de las informaciones de la Web: el objetivo es recolectar todas las páginas sin elegir entre ellas. Las relaciones entre páginas, aun siendo generadas sobre la base de un criterio sencillo, el sistema de los links, generan fenómenos de emergencia. Entonces la calidad de las informaciones, a pesar de su tipología, está determinada por la capacidad de trazar recorridos, sin preocuparnos por cubrir la «totalidad» de las informaciones. La calidad depende sobre todo de la puesta en marcha del propio punto de vista a lo largo de un recorrido de búsqueda: en suma, son los navegantes, los piratas, los usuarios de la Web quienes determinan y mejoran la calidad de las informaciones, trazando relaciones entre las páginas. La potencia de acumulación de los algoritmos de Google es útil, pero insuficiente.

La evaluación del contenido de las páginas y de sus relaciones está delegada a algoritmos y, sobre todo, a las sociedades que los controlan. El fenómeno Google es el resultado de esta delegación: una potencia aparentemente sin límites, en condición de ofrecernos la posibilidad de encontrar «algo» útil e interesante, entre los muchos recursos contenidos en su patrimonio, que se contrabandea por la «totalidad» de la Web. Sin embargo, se silencian completamente las limitaciones de estas brillantes ofertas: lo que falta del patrimonio o lo que está solo parcialmente y, principalmente, todo lo que ha sido descartado.

Sigue siendo de urgente actualidad un delicado problema ético y político acerca de la gestión y el control de las informaciones: ¿qué sujeto puede garantizar la honradez de una empresa, por «buena» que pueda ser, cuyo objetivo primario es el lucro?

Si, por una parte, ingentes recursos económicos y una optima infraestructura tecnológica pueden mejorar continuamente el almacenamiento y la recuperación de datos, las cuestiones políticas derivadas de la acumulación de informaciones no son de ninguna forma superables, pues Google representa de hecho una inédita concentración de informaciones, más o menos reservadas, fuente en todo caso de un enorme poder sin transparencia alguna. Por supuesto, ninguna ley sobre privacidad podrá mejorar esta situación, tampoco la creación de instituciones nacionales o internacionales ad hoc para el control de las informaciones personales y de los datos sensibles. Una respuesta sensata a la cuestión de la confidencialidad de las informaciones solo puede llegar de una responsabilidad y conciencia mayores por parte de los individuos que crean la Red, mediante decisiones de autoformación; un camino difícil y trabajoso, pero el único que puede dar frutos.

El mito de la búsqueda instantánea

Aclarado que el enorme patrimonio de Google nunca podrá corresponderse con la totalidad de la Red, es ingenuo —o, por el contrario, audaz— proponerse como una interfaz «instantánea» entre las peticiones de búsqueda de los usuarios y el resultado supuestamente «exacto».

Cada vez que se efectúa una navegación visitando páginas web se siguen unos links que crean recorridos analizables según los modelos matemáticos de la teoría de los grafos. La Web está de hecho compuesta por nodos (las páginas y los sitios de Internet) y por arcos (los links que los unen).

Por mucho que puedan crecer las dimensiones de la Web, las directrices propuestas por los motores de búsqueda serán siempre útiles para llevarnos al objetivo preciso. Mediante la aplicación de criterios de eficacia/eficiencia, el recorrido trazado por una pregunta en un motor de búsqueda será optimizado, es decir, el número de nodos visitados será muy bajo y el tiempo para hallar el resultado será casi instantáneo. Google llega a auspiciar un solo paso, hasta el punto de ofrecer el botón «voy a tener suerte» en su página principal.

Esta optimización limita la búsqueda a un esquema secuencial articulado en tres puntos: *usuario-algoritmo-objetivo*. Con el tiempo esta dinámica lleva a una postura de pasividad digital, «esperamos» que se nos ofrezcan unos datos para poder «elegir» entre ellos.

Además esta eficacia/eficiencia paradójicamente no se fundamenta en la ampliación del cauce de datos en los que efectuar las búsquedas, sino en la reducción del acceso al mismo patrimonio informativo, pues cada recorrido propuesto por el motor de búsqueda no será nunca trazado en virtud del estado de la Red en el momento en que navegamos, sino calculado en base al patrimonio almacenado y personalizado sobre el usuario a base de filtros algorítmicos y *cookies*.

El acceso a las informaciones propuesto por Google es rápido, mejor, en sus intenciones tiende a ser inmediato, a anular la dimensión del tiempo, a pesar de la enorme extensión de las bases de datos consultadas. La mediación de la tecnología (interfaces, algoritmos, búsquedas preconfeccionadas, etc.) permite esta «anulación» temporal y produce un acceso prácticamente instantáneo[4]. La rapidez en alcanzar el resultado va en detrimento de la calidad de la búsqueda. Cualquiera que haya intentado llevar a cabo una búsqueda sabe bien que el tiempo empleado es un factor determinante

4 La satisfacción de los usuarios con los resultados obtenidos por este método «inmediato» es otra cosa muy diferente: véase a este respecto Jean Véronis, *A Comparative Study of Six Search Engines*, Université de Provence, http://www.up.univ-mrs.fr/veronis, http://aixtal.blogspot.com.

en la composición de la experiencia: crear un recorrido propio, tomar diferentes decisiones según el momento genera sentido de participación y satisfacción. Google nos permite «localizar» en el espacio (en su espacio multidimensional) lo que queremos; pero por muy breve que pueda ser el tiempo de espera del resultado, siempre estamos asumiendo una postura pasiva frente a un oráculo tecnológico.

En el marco de la exploración activa, no obstante, el objetivo no es acceder a los datos, sino realizar una navegación rica y diversificada, construir recorridos complejos. El concepto de eficiencia decae. Cuanto mayor sea el número de nodos que visitamos y la complejidad del retículo que construyamos, tanto mayor será la oportunidad de ampliar nuestros campos de interés, las posibilidades de operar elecciones significativas, de refinar nuestras búsquedas. Este acercamiento permite un enriquecimiento cognitivo más allá de la prestación inmediata. Por ejemplo, cuando visitamos los links propuestos por un sitio de nuestro interés y seguimos los sitios señalados como amigos, creamos un recorrido único; a lo mejor aprovechando un marcador nuestro. A nivel visual, un proceso como éste se opone a la linealidad de la secuencia *usuario-algoritmo-objetivo* y crea, por el contrario, un recorrido lleno de bifurcaciones, ramificaciones, saltos, giros retorcidos, respondiendo a un deseo cognitivo no lineal[5].

En definitiva, los motores de búsqueda son excelentes herramientas en cuanto a la satisfacción de los aspectos cuantitativos de la búsqueda, cuando nos movemos dentro de recursos ya totalmente estructurados, como diccionarios, enciclopedias, etc. La cantidad es directamente proporcional a la potencia de cálculo y de acumulación; la extensión de Google es decididamente superior a la de cualquier otro competidor, pero para mantener esa distancia necesita de

5 El texto básico para acercarnos a las dinámicas cognitivas no lineales sigue siendo el de Humberto R. Maturana y Francisco J. Varela, *Autopoiesis y Cognición. La realización de lo viviente*, Lumen, Buenos Aires, 2004 (ed. or., *Autopiesis and cognition. The Realization of Living*).

una expansión constante, en recursos, algoritmos, máquinas, usuarios, etc.

La máquina de la acumulación se alimenta de la búsqueda de la eficiencia y de la excelencia a nivel de prestaciones, prescindiendo de la evaluación de los contextos de búsqueda: el interés principal de un motor de búsqueda es el incremento de su propio patrimonio y, en consecuencia, de su propia oferta. Una lógica exclusivamente cuantitativa se limita a la idea de que cuantas más sean las informaciones disponibles, tanto mayor será la posibilidad de acceder a informaciones coherentes con la intención de búsqueda: tesis que de nuevo confirma el esquema lineal común al negocio de los motores de búsqueda.

Por lo contrario, la calidad no reside necesariamente en la potencia tecnoeconómica. No es ni mucho menos evidente que la amplitud de los almacenes de información disponibles se corresponda con la pertinencia de los resultados ofrecidos: la emergencia del mejor recorrido no se prevé, no se calcula, sino que se construye paso a paso.

Detrás del velo del mito

Los valores del posicionamiento del ranking de Google no responden a ningún criterio transparente de evaluación; sin embargo, las respuestas ofrecidas son en su mayoría exhaustivas o, mejor dicho, para nosotros es imposible evaluar si algo se le ha escapado a su spider, a no ser que seamos expertos de un determinado sector y conozcamos un recurso que no sale indexado.

La difusión capilar de los instrumentos de búsqueda de Google lo convierten en un estándar de facto. La «caja vacía» (*blank box*) donde introducimos las palabras clave (*keywords*) de nuestras búsquedas es para el usuario una especie de *Weltanschaung* o cosmovisión que propaga una «visión del mundo» bien precisa, puesto que encarna la idea de «ser-

vicio total»: el motor de búsqueda es capaz de contestar a cualquier pregunta, de satisfacer todas las exigencias frente al instrumento Internet.

En términos epistemológicos, la *blank box* representa un modelo cognitivo de organización del conocimiento: pedimos al espacio en blanco la respuesta para todas nuestras intenciones de búsqueda, bien sea cuando necesitamos documentos, profundizaciones, informaciones, o cuando simplemente deseamos «navegar». La actividad de búsqueda se identifica por completo con el objeto que ofrece el servicio, Google, del cual tenemos una percepción redundante.

La costumbre de utilizar las herramientas se transforma en comportamiento, en coacción a repetir: es muy difícil que de motu propio los usuarios intenten satisfacer su «necesidad de inputs» con otras modalidades, habiéndose aficionado a la confortable eficacia/eficiencia de la *blank box*.

Estar en movimiento en la Red y necesitar, por tanto, de interfaces de acceso, de herramientas para encontrar informaciones y construir recorridos, es una actividad diferenciada y profundamente contextual. La actividad de búsqueda no es nada homogénea y no se puede reducir a la utilización de la *blank box*. Lo que pedimos y deseamos no depende solo de una necesidad expresable en términos analíticos de información cuantitativa, sino también de cómo nos relacionamos con la búsqueda, del contexto de la misma, de nuestro bagaje cultural y de la actitud frente a la novedad, a nuevos territorios, a la diversidad. Es imposible agotar las peticiones de información por medio de una solución unívoca.

Como la indexación de las páginas es parcial, se trata pues de una selección basada en el ranking, lo que Google nos propone es, simplemente, la posibilidad de encontrar «algo» útil e interesante para nosotros entre las muchas cosas que están almacenadas en su patrimonio de interés. Un deseo de búsqueda, no obstante, supone en última instancia la voluntad de encontrar, más bien descubrir, «todo lo que no se sabe pero que es posible saber». Entonces el gigante se nos aparece tal y como es: enorme, extenso, ramificado, pero no necesariamente apto para nuestras búsquedas.

Modelos de búsqueda

La ambigüedad sugerida por los motores de búsqueda según los cuales «buscamos en un ambiente infinito», en lugar de en un mundo cerrado localizado sobre la base de nuestros intereses, depende del solapamiento formal entre el nivel de la interfaz[6] y el de la organización. En este contexto específico, la interfaz es el elemento tecnológico que garantiza el acceso a las informaciones y la ejecución de la actividad de búsqueda; la organización es, por el contrario, la arquitectura, el modelo tecnológico con el que las informaciones son ordenadas y almacenadas. Los dos niveles se influencian entre ellos, de hecho las elecciones en términos de organización determinan la utilización de interfaces específicas y, a su vez, las informaciones visualizadas por medio de las interfaces reflejan la forma en la que se conservan.

El problema de este solapamiento reside en que los datos se presentan de manera unívoca y separados entre ellos. De esta forma, el usuario de Google se mueve de forma lineal en la lista de los resultados del ranking; para pasar de un resultado a otro debe volver al listado originario, sin conexiones transversales al nivel de la interfaz.

De hecho, con los motores de búsqueda se encuentran informaciones sin tener en consideración el recorrido realizado. La interfaz que determina nuestra interacción es la *blank box,* adonde se introducen las *queries* o formulaciones para dar comienzo a la búsqueda. En este primer nivel de acceso todas las informaciones se encuentran en el mismo plano y

6 La interfaz es el punto, el área, el territorio que pone en conexión elementos cualitativamente diferentes; en el caso de los motores de búsqueda, la interfaz permite la relación entre el sujeto que interroga y el cauce de informaciones.

son homogéneas, pero al mismo tiempo están separadas y fragmentadas, posibilitando un listado de resultados ordenados por un algoritmo en base a su pertinencia.

Pero en las búsquedas diarias los propios resultados pueden estar entrelazados entre ellos de distintas maneras, sin volver necesariamente al listado ordenado de los resultados, y no hay solo un resultado correcto; es más, una búsqueda que no sea dirigida a datos estructurados en enciclopedias, vocabularios y similares (que también cambian con el tiempo), podría no tener solución alguna, sino necesitar de esfuerzos creativos, «mezclas» y recombinaciones.

Cuando se produce la identidad formal entre el nivel de la interfaz y el de la organización, se obtiene un modelo reduccionista; en el caso de Google nos enfrentamos a la representación de una potencia de búsqueda infinita, o sea, que la actividad de búsqueda se confunde con uno de los instrumentos para realizarla.

Por ejemplo, si buscamos «Penna» («bolígrafo» en italiano) a través de la interfaz de Google.com, encontramos, entre más de seis millones setecientos mil resultados: «Sandro Penna», el poeta; «Penna d'autore», el concurso literario; el «municipio de Penna San Giovanni», en la provincia de Macerata y más cosas. No encontramos entre los primeros resultados (*top ranking*) ni recursos referidos a las «*penne* (plumas) de las aves», ni a las «plumas para escribir», dos conceptos muy lejanos entre sí, sobre los que queríamos investigar eventuales relaciones. Una perspectiva más amplia acerca de la búsqueda de informaciones, que tenga en cuenta críticamente el potencial cognitivo comprendido en todo proceso informativo, nos fuerza a considerar la función acceso/búsqueda no como un proceso de localización, sino de exploración y creación. El acento se desplaza por lo tanto de la epistemología a la ontología: no solo conocer unas informaciones, sino tomar conciencia del rol de creadores de informaciones[7]. Los motores

7 Explorar las redes significa crear las redes: en este sentido, ampliando el perfil analítico de Maturana y Varela, podemos hablar de «autopoiésis de las redes», redes que se construyen a través de nuestro hacer.

de búsqueda intervienen en el acceso pero no ofrecen ningún soporte para la exploración, pues se limitan a intervenir en el primer nivel de presentación de las informaciones.

La navegación es el momento de dinamismo real en el proceso de relacionar los objetos digitales, que pueden así expresar al máximo nivel sus potencialidades comunicativas y heurísticas. Se aprende haciendo, y se cambia mientras se va aprendiendo, mientras se está involucrado en la exploración.

Existe una distinción profunda entre buscar y encontrar. Google nos hace encontrar las cosas, produce satisfacción, sensación de acumulación. Pero podría no ser tan interesante el hecho de «encontrar» como el acto mismo de «buscar»; quizás sea interesante no encontrar nada, pues si no se consigue encontrar quiere decir que estamos sumergidos en la búsqueda[8].

El motor de búsqueda es un modelo instrumental a través del cual se ordena la información: en lugar de eso, quizás sería más interesante y loable imaginar modelos para combinar las informaciones creando conocimiento.

8 :-).

7

Tecnocracia

El análisis del fenómeno Google nos abre un panorama variado, en el que la economía de la búsqueda resulta ser solo una pieza de un mosaico mucho más amplio y complejo. Por otra parte, si nos remitimos a las declaraciones de Eric Schmidt, en Mountain View están poniendo las bases para la construcción de una empresa global de las tecnologías de la información, una «empresa de cien mil millones de dólares»: ciertamente, no se trata de un simple motor de búsqueda. Se trata en efecto de un sistema invasivo de gestión de los conocimientos, del que hemos ilustrado las fases y los métodos más significativos: estrategias que conjugan marketing agresivo y meditada gestión de la propia imagen; propagación de interfaces altamente configurables y a la vez siempre reconocibles; creación de contenidos estándar descentralizada en usuarios y desarrolladores; cooptación de metodologías de desarrollo cooperativo típicas del open source y del free software; utilización de sistemas avanzados de recogida y almacenamiento de datos; sistemas de localización de las informaciones vinculados a técnicas de «perfilación» explícitas e implícitas; personalización de la publicidad.

Google es el dominio tecnológico derivado de la búsqueda científica que se hace instrumento de gestión del conocimiento, expresión directa de la tecnocracia.

Tecnocracia: los expertos de la ciencia

Los expertos han hallado en el control y en la manipulación de la tecnología un arma sin parangón para imponer sus propios intereses personales, para mantener el poder o adquirir mayores privilegios. El mecanismo es simple: la tecnología se utiliza como narración garante de la objetividad de la investigación científica, pero también para avalar las decisiones de los politicuchos que gobiernan y, en general, de cualquier «autoridad» que tenga acceso al oráculo tecnológico.

La aplicación de la investigación científica bajo forma de tecnología es invasiva, y la realidad se interpreta constantemente conforme a este paradigma. La curiosidad y el deseo de saber que mueven la investigación científica se encuentran castrados por los criterios de provecho siniestro que regulan las financiaciones públicas y privadas. Si la investigación no produce artefactos tecnológicos capaces de generar ingresos inmediatos, entonces no resulta interesante. El discurso del poder se convierte entonces en tecnocracia, justo lo contrario de la participación comunitaria, de la autogestión, de la discusión y de la mediación entre individuos. Hacer pasar la tecnocracia de Google por democracia directa es un juego lingüístico que nos da la ilusión de participar en una gran democracia electrónica, carente de toda sustancia. Cierto, podemos publicar lo que queremos en Internet, y Google nos indexará. Pero no podemos permitirnos, siendo «profanos» y «heréticos», observar que la estrategia de la acumulación de Google está en perfecta sintonía con el alocado sistema de la economía de mercado basada en el crecimiento infinito: de hecho, no somos economistas licenciados en la London School, ni empresarios de éxito, es decir, no somos expertos y, por lo tanto, no tenemos ninguna autoridad. Sin embargo, el buen sentido común de orwelliana memoria es más que suficiente para entender que este crecimiento sin fin y sin objetivo es la manifestación de una voluntad de potencia tecnológica interesada en los individuos solo en tanto que potenciales consumidores.

Ya sabemos por qué el PageRank(™), que como hemos visto no es solo un algoritmo, se convierte en una vestimenta cultural con la que Google nos propone analizar cada cosa: en resumen, asistimos a una extensión forzosa del método de revisión por pares, válido en el ámbito científico, hacia la totalidad del conocimiento humano.

Las autoridades tradicionales, religiosas y políticas, han alcanzado el punto mínimo de credibilidad; pero su escasa influencia sobre la realidad, en lugar de favorecer la proliferación de espacios de autonomía, ha llevado a una situación surrealista en la que ninguna aserción válida puede ser completada sin el apoyo de una autoridad tecnológica. La autoridad de las máquinas, en la mayoría de los casos, es una respuesta extraída de una base de datos, servida por los sacerdotes de la tecnología, los «expertos», al rico y acomodado pueblo de los *prosumer*. El relativismo extremo se hace realidad en los métodos para extraer «verdades» de los datos disponibles, casi sin límites, que se corresponden con el número de algoritmos y filtros que se pueden aplicar. Una verdad oportuna para cada búsqueda se corresponde con un producto personalizado para cada consumidor evolucionado.

Frente a este cierre en la creación, gestión y aplicación (en nuestra propia piel) de los conocimientos, parece que quedan solo dos posibilidades: rechazar la cultura científica como si fuese la causa de todo mal; o, por el contrario, aceptar con entusiasmo fideístico toda «innovación» tecnológica. Sin embargo, entre los dos extremos de la tecnofobia y de la tecnofilia, es posible proponer la curiosidad de la ética hacker, es decir, compartir los saberes, la actitud crítica hacia las «verdades», el análisis cuidadoso de las fuentes, insistiendo en el recorrido de la libre circulación de los saberes.

En este contexto, la formación es un nodo central, pero no hay salidas para una educación científica difusa. La estructura de la formación tanto en Europa como en Norteamérica gira alrededor de la sola producción de especialistas. Actualmente no hay ningún modelo pedagógico que responda a la

exigencia de un saber científico de tipo *amateur*, ni siquiera en países de tradición no occidental como Brasil o la India que producen investigación científica de alto nivel y tecnología punta de bajo coste gracias a una competencia despiadada. Un saber científico no académico ni empresarial, difuso y autogestionado, no se admite ni siquiera en fase de discusión: y sin embargo sería indispensable para la formación de competencias de base y para la evaluación de las innovaciones tecnológicas que nos conciernen a todos. Especialmente, la noción global de «educación científica» debería de ser reconstruida para responder a la necesidad difusa de rudimentos prácticos para enfrentarnos a la onda larga de la revolución tecnológica.

La afirmación de la informática como principal motor de la innovación tecnológica abre, en efecto, nuevos escenarios: la informática no es simplemente una técnica para gestionar la información de forma automática, sino que posee una lógica propia, o sea, que trabaja y modifica continuamente sus mismos cimientos. Es física teórica y experimental a la vez: estudia la formalización del lenguaje (formaliza, pues, el conocimiento), la aplica a los componentes físicos de la electrónica y saca de ésta lenguajes posteriores que, a su vez, influencian las teorías del conocimiento; se basa en una tipología de recursos totalmente particular.

En las ciencias clásicas se observan fenómenos estables: la física, por ejemplo, elabora datos naturales y sobre ellos crea teorías; en la informática, por el contrario, los fenómenos elaborados en la teoría son artificiales, cambian física y conceptualmente a medida que los progresos técnicos y experimentales permiten pulirlos: un software elaborado en un ordenador de hace diez años será estructuralmente distinto de uno más reciente. Lo que ayer era verdad, sabemos ya hoy que no será verdadero mañana, cuando tendremos máquinas más potentes para hacer cosas nuevas, porque hablamos de un mundo vivo y por ello constantemente en devenir.

Milagros de la tecnología: de las opiniones subjetivas a la verdad objetiva

El buen Google, en cuanto gigantesca base de datos, interviene en este panorama y sostiene que somos parte de una gran e inédita «democracia electrónica global»; sostiene que los resultados del PageRank(™) son correctos porque son expresión de la democracia directa de los links evaluados por sus algoritmos; nos devuelve pues, en cierto sentido, nuestro legítimo derecho a dar «nuestra opinión».

La popularidad no puede ser considerada índice de «calidad objetiva» desde un punto de vista epistemológico, ya que en este caso el concepto mismo de objetividad se basa en un malentendido, es decir, en la convicción de que muchas ideas subjetivas (las «opiniones» expresadas bajo forma de link/voto) se convierten por encantamiento en su exacto contrario, en objetividad revelada, apenas superan un determinado número y se convierten en mayoría. De este modo, el ranking se transforma en expresión de calidad porque es fruto explícito de una tecnología de manipulación de la información.

¿Pero cómo es posible que la cantidad se convierta en calidad? Se asume subrepticiamente que la mediación técnica del algoritmo es garante de «objetividad», y se asocia a la objetividad la calificación de «bueno», más aún, de «mejor», «verdadero». Todo tiene que ser rápido o, mejor dicho, inmediato y transparente, gracias a la eliminación del factor tiempo en la búsqueda y al estudio ergonómico de la interfaz.

El mecanismo de creación del consenso que Google considera manifestación de la «democracia directa» de los usuarios/votantes no convence principalmente por dos razones: por un lado, asume que la mayoría tiene siempre razón; por el otro, implica que las opiniones de la mayoría, para ser disfrutadas por los usuarios, tienen necesariamente que pasar a través de una mediación técnica que nunca es explicitada como tal.

La dicotomía objetivo/subjetivo, sobrepuesta a la de la verdad/opinión, es del todo inadecuada en el mundo de las redes. Para ser precisos, la ciencia siempre ha creado híbridos naturales-culturales, es decir, ha inventado técnicas y promovido tecnologías. Por un lado, la observación y la experimentación tienen como campo de acción la «naturaleza», y en este sentido son «objetivas»; por el otro, sus resultados son altamente subjetivos. La ciencia, de hecho, está sometida a la voluntad y a la percepción político-social, en cuanto mediada por el lenguaje (como mínimo, el lenguaje de la comunicación científica) y en cuanto fuente de poder (qué mejor ejemplo que la bomba atómica).

La tecnología que implementa las redes es la aplicación contemporánea del método científico para crear «híbridos naturales-culturales», los enésimos objetos científicos que luego son propuestos cada vez más como testigos de la realidad en lugar de los humanos[1]. Ya tiene más valor la respuesta del algoritmo de PageRank(™), un híbrido tecnológico, que la de cualquiera persona, y a menudo el resultado obtenido en Google incluso tiene más valor que la opinión de un experto del sector, aunque solo sea por el hecho de que PageRank(™) está siempre al alcance de un clic y un experto no.

Como se ha subrayado con anterioridad, desde un punto de vista materialista, después de todo, Internet es natural, en el sentido de material, es decir, que está constituida por máquinas mecánicas y electrónicas y, a la vez, es cultural, porque no existiría sin el sentido que la cultura le atribuye, constituida por interacciones significativas de los seres humanos o, mejor, de las máquinas biológicas entre sí y con las máquinas electrónicas. El carácter híbrido de las redes es una consecuencia obvia del carácter híbrido de la tecnología a secas.

1 A propósito de la visión de la tecnología como «puente» de hibridación entre naturaleza y cultura (extremidad teórica de la esquizofrenia del sujeto moderno), véase la obra de Bruno Latour, y en particular *Non siamo mai stati moderni – Saggio di antropología simétrica*, Elèuthera, Milán, 1994. Un punto de vista parecido es el de Gilbert Simondon, que elabora un discurso sobre la «técnica transindividual», un espacio mestizo constituido por lo individual y lo colectivo al mismo tiempo.

Otro punto de vista posible sobre la cuestión subjetividad/objetividad tiene que ver con el modelo de decisión: cómo se decide lo que es relevante. En un contexto relativista, es fácil concebir una información procedente de un sitio (sea un blog, Google o un circuito de información oficial) como «objetiva» si el juicio es el resultado de un posicionamiento claro, de una clave de lectura hecha pública, de la visión de nuestro pequeño mundo localizado. Una red de confianza, es decir, un grupo de personas que intercambia informaciones, pareceres y, en general, conocimiento, puede comunicar fácilmente el propio modo de trabajar, la propia jerarquía interna eventual y las modalidades para adherirse a aquella red o proyecto específico. Cada vez que se consulten los resultados de esta red de confianza, éstos podrán ser leídos como «objetivos», en el sentido de verdaderos de por sí, relevantes para la propia experiencia, precisamente porque son fruto de tantas subjetividades distintas y de una fuerte interacción entre ellas. Si nos sentimos en sintonía con esta comunidad se podrá considerar interesante la información hallada o descartarla en función de otras redes de confianza.

Según este enfoque, si Google comunicara de forma pública el mecanismo propio de decisión y los usuarios de Internet estuviesen en condiciones de comprenderlo se podría superar fácilmente el problema objetivo/subjetivo, llegando a identificarse poco a poco, búsqueda tras búsqueda, con la red que más nos satisface, y a influenciarla directamente a fin de que nos refleje a nosotros, con nuestro gustos, nuestras ideas y nuestras idiosincrasias.

Público y privado

El PageRank(™) ilustra también otra dicotomía: la que se da entre lo público y lo privado. En los hechos, todo lo que pasa

a través de Google se publica, se hace público: ¿a quién no le ha pasado de encontrar e-mails personales en los resultados del ranking, quizás porque habían sido enviados a mailing-lists públicas? Ya que una cantidad cada vez mayor de informaciones personales pasa a través de Google, y en general a través de las tecnologías informáticas, la posibilidad de reencontrar mañana las propias llamadas (tal vez realizada vía VoIP), archivadas y disponibles a través de un motor de búsqueda, parece todo menos remota. En suma, las tecnologías llegan a hibridar incluso las esferas de lo público y de lo privado; por otra parte, conectarse a Internet significa abrir una puerta al mundo, y a través de esta ventana seguramente el mundo entra en nuestra vida.

Existen ya algunas redes cuyas prácticas desautorizan de hecho la quimera de una información objetiva, ejerciendo una elección precisa, deliberada y del todo subjetiva entre lo que intentan hacer público y lo que, en cambio, concierne a la esfera privada. Este fenómeno toma cuerpo en su forma más extensa cuando el motor de búsqueda se revela incapaz de satisfacer una instancia de búsqueda cuya calidad (específica) va más allá de la disponibilidad cuantitativa y de la estructura técnica propuesta.

El ejemplo más conocido es la búsqueda trámite en las redes *peer-to-peer* (P2P)[2] como eMule u otros parecidos. La

2 Literalmente, P2P significa «de par a par»; en general, se trata de un modelo de comunicación en el que cada una de las partes tiene las mismas funciones y cada una de las partes puede empezar la sesión de comunicación, frente a otros modelos como el server/client o el master/slave. A menudo la comunicación P2P se ve implementada ofreciendo a cada uno de los nodos de comunicación las funcionalidades de server y client. En el lenguaje corriente, el término *peer-to-peer* se utiliza para describir las aplicaciones mediante las que los usuarios pueden, a través de Internet, intercambiar directamente archivos con otros usuarios. En particular, por lo que respecta a Internet, P2P es un tipo de network «transitorio» que permite a un grupo de personas, que utilizan el mismo programa, conectarse y acceder directamente a los recursos compartidos. Napster hasta hace algunos años y ahora Emule, Lopster, etc., son ejemplos de estos softwares.

cantidad de datos que se pueden buscar en estas redes se corresponde con los datos compartidos por los usuarios y varía en el tiempo de forma no periódica; por esta razón se les define como redes «transientes»[3]. En efecto, en estos casos el usuario realiza libremente una elección entre el material que desea hacer público y lo que, por otra parte, se quedará sellado en la esfera privada; también porque una regla básica del intercambio «entre pares» o «iguales» es el hecho de ofrecer compartir los propios materiales, que se pueden buscar en la Red, para conseguir a cambio otros materiales. En sí el P2P es una práctica legal; en el límite, los datos intercambiados pueden estar protegidos por copyright (audio, vídeo, textos, etc.): solo en ese caso se entra en la ilegalidad. La difusión planetaria de estas prácticas pone de manifiesto la naturaleza liberticida de la legislación vigente en materia de propiedad intelectual y de patentabilidad del conocimiento: actualmente constituye un delito penal difundir informaciones protegidas por el copyright, es decir, que una buena cuarta parte de la población italiana debería acabar en prisión por haber descargado al menos un archivo mp3 ilegalmente. Además, parece cada vez más urgente la decisión individual sobre la frontera entre lo público y lo privado. Se abre un amplio abanico de posibilidades que se mueven entre la decisión ponderada de favorecer e impulsar el intercambio de informaciones más allá de cualquier consideración legal y el simple recurso a la piratería, aprovechando lo que se encuentra en la Red.

Piratería difusa no significa ciertamente estar a las puertas de la revolución popular, también porque a menudo se trata de una piratería semiinconsciente, no de una asunción

3 El término «transiente» sinónimo técnico de «transitorio», procede de la astrofísica y de la acústica y se refiere a una fuente cuya radiación varía en el tiempo y de forma no periódica. Las redes «transientes» son transitorias desde el punto de vista de los flujos informacionales que las componen; en el caso de las redes P2P, la «transientidad» depende de la variación de la cantidad de informaciones compartidas por cada individuo.

de responsabilidad por parte de los individuos que eligen oponerse al vigente sistema de protección del conocimiento[4]. Más bien las tecnologías digitales transforman la capacidad de reproducción técnica en un problema obsoleto, y la cultura del consumismo empuja a desear mucho más de lo que se puede utilizar. Es casi obvio desear todo lo que conseguimos imaginar, por lo menos a nivel de «informaciones», aun cuando sea evidente que nunca seremos capaces de escuchar ni siquiera una fracción de la música que descargamos de la Red, o de ver incluso solo una porción significativa de las películas que guardamos en nuestro hard disk. Pero la instancia del deseo, por su propia naturaleza ilimitado, unida a la posibilidad técnica, abre serios interrogantes sobre la distribución y el acceso al conocimiento. El carácter tendencialmente gratuito y libre de estos intercambios constituye un reto extraordinario a la primacía de la economía productiva, pero la difusión de las opiniones, por ejemplo bajo forma de blog, pone en crisis incluso la concepción tradicional de los mass media.

El P2P representa de hecho solo el nivel superficial, y más difuso, del intercambio no mediado por una autoridad genéricamente supracomunitaria. Hay muchas otras instancias de búsqueda cualitativa, todas satisfechas por redes de confianza y capaces de recombinar la perspectiva de nuestra orientación en Internet. En algunos casos estos recorridos subjetivos convergen en sistemas de afinidad cultural o profesional, por ejemplo fórums, newsgroups, blogs especializados, o en el disenso explícito de las fuentes oficiales; no solo, a menudo constituyen un ejemplo suficientemente articulado de un modelo alternativo para la organización del conocimiento (*knowledge management*).

4 Y, sin embargo, violar las leyes sobre el copyright es un delito tan increíblemente difuso como grave: ¡un delito penal!, y así se regula en la normativa europea conocida como EUCD, modelada sobre la estadounidense DMCA, del 2003. Véase http://www.softwarelibero.it/progetti/eucd/annuncio-campagna.shtml.

Es por lo tanto fácil plantear la hipótesis de una (tal vez lenta) evolución del circuito de conocimiento relacionado con el instrumento del blog hacia instrumentos P2P. Intercambio de conocimiento a través de programas desligados de la publicación web y basados en redes dinámicas en las que sea posible compartir flujos de informaciones y archivos entre los usuarios conectados y, por lo tanto, realmente pertenecientes al propio circuito de amigos. Esto permitiría hacer que la división entre público/privado no sea algo estático, sino dinámico y difuminado, de manera que cada persona sea libre de compartir informaciones escogiendo un nivel de accesibilidad con un valor intermedio entre lo público y lo privado[5].

Líneas de fuga: media independientes, criptografía, blogs, FOAF

El fenómeno de los blogs[109], sitios web que contienen reflexiones personales y links del autor y comentarios de los lectores, ha dado lugar a lo que ya se califica como *blogsfera*: el informe de Technocrati, publicado en octubre de 2008, calculaba que en ese momento existían 150 millones de blogs en activo, duplicando el número de los existentes en 2007, año en que la blogosfera aumentaba en más de cien mil blogs al día, duplicando esa cantidad cada seis-siete meses[6].

Desde el punto de vista matemático, los blogs respetan la ley de potencia y se distribuyen según la característica «larga cola»: pocos centenares de blogs tienen un número muy elevado de links (cerca de cuatro mil constituyen la flor y nata de los blogs), la mayor parte (millones) tiene poquísimos links. En este sentido, como hemos visto en el Capítulo 2, la blogosfera pertenece al marco de la economía de la búsqueda. El

5 Con sus aspectos positivos y negativos, la expansión de un recurso como Facebook en los últimos años parece caminar en esta dirección. [N. del. E.]

6 Véase el informe *La blogosfera e i media*, www.casaleggioassociati.it.

blog permite compartir el propio punto de vista, declaradamente subjetivo y personal, sin ningún filtro; el número de links a un blog representa su popularidad y, por lo tanto, su autoridad, que puede alcanzar y superar la de los periódicos y otros media tradicionales, generando fenómenos de influencia sobre la opinión pública. Fiabilidad, confianza y reputación están asociadas al autor particular del blog: a medida que crece la popularidad se hace difícil mentir sin ser penalizado inmediatamente, es decir, sufrir deserciones y acabar cayendo en el olvido[7].

La autoridad del blog de Beppe Grillo, el único italiano que aparece entre los primeros 100 blogs del mundo por número de links, es mayor que la de los blogs de *La Repubblica* e *Il Corriere della Sera*. El sujeto Beppe Grillo se expone a nivel personal y no pretende vender verdades: expone su punto de vista. En este sentido, los blogs crean redes compartidas autogestionadas; no solo, a veces pueden convertirse en las únicas fuentes de información no alineadas. Es el caso del *blogger* iraquí Salam Pax (alias *Salam al-Janabi*) durante la segunda Guerra del Golfo[8].

La novedad más importante ligada a la información en los blogs es la de la agregación automática de fuentes distintas a través de los *feed* RSS (literalmente: «provisiones de RSS»), el estándar de facto para la exportación de los contenidos de la Web. En pocas palabras, se trata de un método automático para deslizarse rápidamente en sitios web, y encontrar y exportar las informaciones que interesan. La popularidad de

7 Existen varios métodos para maquillar los datos relativos a los links a un blog, por ejemplo las *link farms* o «granjas de enlace» (grupos extensos de páginas web creadas para aumentar la popularidad de una página), o los *fake blogs*, o sea, blogs creados de forma automática por programas, igual aprovechando recopiladores de noticias tipo RSS. Google intenta desde hace tiempo solucionar este problema que genera el fenómeno de los clics fraudulentos y que afecta, al menos en parte, el mecanismo publicitario.

8 La vida en Bagdad bajo los bombardeos en el blog de Salam Pax: http://dear_raed.blogspot.com/.

los blogs es tal vez una de las razones principales del éxito de RSS: millares de *weblogs* empiezan a producir contenidos en RSS y así proliferan sitios que recogen una selección de *posts* de los blogs más seguidos (los llamados *blog aggregators*) y programas para usar los contenidos de un blog directamente en el propio terminal (*RSS reader*). De esta manera es posible buscar contenidos en la blogosfera sin pasar por Google.

La posibilidad de poder recibir cómodamente de forma automática los últimos artículos escritos sobre los argumentos que el usuario considera más interesantes en el propio ordenador es desde luego una innovación destinada a tener enormes repercusiones en el uso de la Web. Más en general, ya que el RSS es un primer paso para disponer de toda información en un formato que por su propia naturaleza se puede compartir fácilmente, transformar y ampliar, hace que una información sea accesible a todo soporte informático, sea un sitio, un programa instalado en el propio ordenador, móvil o cualquier otro instrumento tecnológico.

Sin embargo, existen muchas otras posibilidades para buscar informaciones. Si Google se presenta como un medio objetivo y público, Indymedia.org[9], por ejemplo, se define como «una red de medios de comunicación gestionados colectivamente para una narración radical, objetiva y apasionada de la verdad», donde, sin embargo, existe un grupo de personas que actúa según una política pública muy específica: en la columna a la derecha de las páginas, llamada *newswire*, cualquiera puede publicar. No se censura nada, pero «los mensajes patentemente racistas, sexistas, fascistas se ocultan, sin ser eliminados». Indymedia es esencialmente un instrumento para usuarios activos en la creación de la información y para contribuir a la formación del conocimiento y de verdades compartidas: la autoridad que «crea» las verdades es difusa y participada, es una autoridad ética (colectivo) y no matemática (algoritmo).

9 www.indymedia.org; http://www.indymedia.org/es/index.shtml.

Poner en discusión las fuentes oficiales demostrando que es posible hacer información de un modo creíble, descentralizado e independiente es uno de los objetivos de Indymedia.org y de decenas de networks surgidos alrededor de temas específicos[10].

¿Publicar las informaciones o mantenerlas privadas? Una respuesta interesante procede del proyecto Freenet.org[11], una red P2P descentralizada, creada para resistir a la censura, que explota los recursos (ancho de banda, espacio en disco) de sus usuarios para permitir la publicación y el disfrute de cualquier tipo de información. Freenet.org se ha constituido pensando en el anonimato y en la seguridad, no en la velocidad de transmisión. Las naciones y autoridades de distinta naturaleza que censuran las comunicaciones comparten todas una característica en común: existen unos controladores que deciden qué cortar y qué mantener, qué se considera ofensivo y qué se considera inocuo. Freenet.org es una red que elimina a cualquiera la posibilidad de imponer su escala de valores: nadie puede decidir lo que es aceptable, la tecnología se pone al servicio de la libertad absoluta de expresión.

En las redes no anónimas, si queremos que nuestras informaciones lleguen exclusivamente a destinatarios preestablecidos, resulta útil la criptografía pesada; y, por otra parte, es una práctica fundamental para garantizar un mínimo de privacidad en las comunicaciones digitales. Por ejemplo, nuestros correos electrónicos encriptados aunque se puedan buscar mediante los instrumentos de búsqueda, no son legibles[12].

10 Por ejemplo: http://www.cyberjournalist.net; http://alternet.org/; www.warnews.it; http://zmag.org/; http://slashdot.org; http://wikipedia.org; http://digg.com, etc.

11 Para entrar a formar parte de la red Freenet: http://freenetproject.org/.

12 La criptografía le da miedo a los gobiernos porque genera espacios privados no inmediatamente accesibles al control policiaco: no por nada la legislación estadounidense consideraba, al menos hasta hace pocos años, «armas» los instrumentos de encriptación pesada. Actualmente, la legislación vigente es muy confusa. Naturalmente la posibilidad de desencriptar un documento depende de la potencia de cálculo que se emplea. No se trata pues de seguridad absoluta, pero sí racionalmente proporcionada a la evolución tecnológica. Así utilizar claves de encriptación asimétrica pesada como el GPC (*Gnu Privacy Guard*) a 204 kb ofrece, en 2006, una protección muy elevada, aunque no total. The Gnu Privacy Guard: http://www.gnupg.org/.

La encriptación se propone usar las redes de forma subjetiva y filtrada, creando espacios privados de intercambio entre particulares en lugar de espacios públicos indexables y aprovechables por cualquiera. No se trata de tener algo que ocultar, sino de preservar espacios privados y de poder decidir autónomamente y libremente qué, cómo y cuándo hacerlo público; nuestras comunicaciones personales, nuestros datos sensibles no serán públicos ni objetos de marketing por el solo hecho de estar transitando en la Red.

Es posible también hacer que todo sea público o, mejor, decidir publicar la visión subjetiva de uno mismo. La propia identidad virtual (álter ego digital) puede ser definida en los mínimos detalles y no dejada en manos de las técnicas de personalización. La propuesta quizás más interesante la representa la comunidad FOAF, acrónimo de *Friend-of-a-Friend*[13], o sea, «amigo de un amigo». El objetivo de este proyecto es crear un conjunto de páginas web legibles automáticamente por las máquinas (*machine-readable*) que describen las personas, sus intereses, las cosas que hacen y sus conexiones recíprocas[14]. De esta forma, el concepto mismo de verdad pasa a un segundo plano, sustituido por el de «confianza»: tiendo a confiar en los amigos de los amigos, formando «redes de confianza» basadas en los gustos, afinidades, pasiones en común. FOAF es un instrumento útil para crear redes sociales, favoreciendo una actitud activa: es preciso «declararse», «describirse», «publicarse»[15].

13 www.foaf-project.org.

14 Más técnicamente, FOAF es un vocabulario EDF/XML para la web semántica con el objetivo de traducir en un lenguaje comprensible a las máquinas (por ejemplo los spider) lo que usualmente está escrito en una página personal sobre el autor de un sitio o de un proyecto.

15 Escribir el propio documento FOAF no es difícil utilizando instru-

Desde el punto de vista de las redes de confianza (*trusting networks*, base de la Web Trust), FOAF puede ser una garantía de transparencia sobre los recorridos que unen nuestros avatares virtuales, favoreciendo un crecimiento de credibilidad y confianza en las relaciones interpersonales desarrolladas en la Red. La Web se utiliza para crear espacios en los que las personas están interconectadas en un gran mapa relacional, que se puede recorrer a partir de cualquiera de sus nodos y sin la necesidad de un instrumento de búsqueda centralizado.

Es posible usar instrumentos como FOAF también para fines declaradamente políticos: por ejemplo, la red Indyvoter, que se define como «un sitio web de *networking* social para activistas políticos», con el objetivos explícito de formar un tipo de gobierno alternativo[16].

Las falsas alternativas binarias, naturaleza/cultura, subjetivo/objetivo, público/privado, verdadero/falso se revelan en su escasez. La informática es constitutivamente dual e impura: teoría y práctica a la vez, crea objetos que modifican nuestra forma de pensar que, una vez utilizados, modifican la propia informática. Los mundos digitales son reales, y en la realidad las visiones en blanco y negro dejan sitio al gris y a la infinita variedad de colores, a los matices de las opiniones, a las diferencias que son riqueza, a lo heterogéneo que no puede ser reconducido a lo único. Los instrumentos para modificar, expandir y ramificar estos espacios de libertad y sus confines están a nuestra disposición: las posibilidades están limitadas solo por nuestra superficialidad y falta de imaginación.

mentos como Foaf-a-matic, http://www.ldodds.com/foaf/foaf-a-matic.it.html.

16 www.indyvoter.org: «La Liga de los Jóvenes Votantes organiza encuentros de votantes fuera de Nueva York. Nuestra misión es involucrar a jóvenes cabreados de 17 a 35 años de edad en el proceso democrático para construir una mayoría de gobierno progresista durante nuestra vida».

Conclusiones

Hemos llegado al final de nuestra exploración, tras haber desvelado muchos secretos, más o menos importantes, de Google.

Como hemos visto se trata de un dispositivo que presume de poder gestionar globalmente el complejo sistema de los conocimientos disponibles en Internet. En efecto vende por verdades objetivas respuestas que son solo recorridos subjetivos filtrados por sus tecnologías de búsqueda. No tenemos que caer en la trampa de una lectura esotérica del fenómeno, estupefactos por la rapidez con la que atiende nuestras solicitudes. Estos portentos, a menudo impregnados de misticismo, son en realidad estrategias que conjugan el uso de sistemas avanzados de recogida y almacenaje de datos con sistemas de localización de las informaciones, relacionados con técnicas de personalización directas e indirectas vinculadas a la perfilación de la publicidad.

Pero no solo el marketing de vanguardia y una oportuna gestión de la comunicación permiten explicar la evangelización googleana: tomemos, como ejemplo, el empleo de colores en la paleta visual del logo; añadamos la difusión de unas interfaces altamente configurables, capaces de conservar la

propia *corporate-identity* en cualquier ámbito, y el plato estará listo: y estará servida también la explotación a todo nivel de la economía relacional puesta en marcha respecto a los usuarios. Y, además, Google se apropia sistemáticamente de las metodologías de desarrollo cooperativo típicas del open source y del free software, abaratando los costes para la implementación de sus propios servicios mientras se hace pasar por defensor de la causa de la libre circulación de los saberes.

Rechazar la hagiografía corriente del «motor de búsqueda perfecto» no significa empezar una campaña de boicot. Por otra parte, el grupo de Ippolita ha usado a menudo Google, incluso para escribir este libro.

Del mismo modo, el empleo frecuente de recursos difusos y accesibles como Wikipedia se ha revelado muy útil a nivel enciclopédico. En efecto, si ya se conoce la cuestión en discusión, se puede comprobar la exactitud de las informaciones de forma independiente, recombinando las partes de la Red de modo laico y desencantado para hablar de la propia Red. El uso crítico de las fuentes depende de la capacidad de los sujetos para aceptar y evaluar la seguridad de las informaciones, no de la bondad intrínseca de las tecnologías digitales.

La informática no es simplemente una técnica para gestionar la información de forma automática, sino que posee una lógica intrínseca, es decir, que trabaja y modifica continuamente sus propios cimientos. Es física, teórica y experimental a la vez: estudia la formalización del lenguaje (formaliza el conocimiento pues), la aplica a los componentes físicos de la electrónica, obteniendo lenguajes posteriores que, a su vez, influencian las teorías del conocimiento; es decir, se fundamenta en una tipología de recursos muy peculiar.

Google desarrolla a fondo esta lógica recursiva como una máquina extraordinaria que se construye mediante el uso por parte de los usuarios. En este sentido, es una máquina «autopoiética», que acumula todas las informaciones de base volcadas cada día en la Red por millones de usuarios (nombres, fotografías, mails, preferencias de búsqueda, participación en fórums, uso de blogs, compilación de cuestionarios y módulos, recorridos de navegación, etc.) para vender publicidad de modo capilar. Los datos de los usuarios se han convertido en un enorme patrimonio económico, social y humano. Un patrimonio a proteger, desde luego, pero también un territorio en el que curiosear, indagar, experimentar.

Google facilita una respuesta fluida a las intenciones de búsqueda de los usuarios, una respuesta articulada en servicios cada vez más sofisticados y personalizables. Pero se trata de una multiplicidad solo aparente, que simplemente se preocupa por difundir una forma de consumismo adecuada a la economía informacional: la personalización en masa de la publicidad y de los productos. El capitalismo de la abundancia de Google practica un cuidadoso registro del imaginario de los productores-consumidores (*prosumers*) a todos los niveles. De hecho, los usuarios facilitan gratuitamente sus datos personales, pero también sugerencias e impresiones sobre el uso de los servicios; los desarrolladores colaboran en la consolidación de los instrumentos «abiertos» puestos a su disposición para difundir los estándares de Google, los cuales quedan bajo el atento control de Mountain View y los empleados de Googleplex y de los demás *datacenters* se identifican plenamente con la filosofía empresarial de la excelencia.

La personalización del imaginario no es sino la última etapa del proceso de colonización capitalista en las redes de lo que hemos llamado onanismo tecnológico. La mentalidad del beneficio se oculta bajo declaraciones a favor de la «libre expresión de los individuos», que luego será aprovechada para vender brillantes e inútiles productos personalizados.

Google promueve su «economía de la búsqueda» como si fuera una nueva ciberdemocracia en la que centenares de millones de personas pueden comunicar directamente y autoorganizarse, sustrayéndose al control estatal e institucional en general a través de sus instrumentos tecnológicos. Este mensaje simplista halla el apoyo convencido de los media y de los intelectuales «democráticos» de todo el globo. Según esta vi-

sión, la naturaleza de Internet sería intrínsecamente democrática: en la Red, no solo los individuos se verían alentados a ocupar el lugar de las instituciones, sino que incluso las propias instituciones devendrían mejores. El entusiasmo tecnocrático llega hasta el punto de presentar la informatización de la administración pública, un proceso conocido como *e-government*, como una especie de gobierno sin ideologías, que apela a la responsabilidad del «ciberciudadano». Esta identidad nueva por fin participa (digitalmente) en primera persona de la emersión de una opinión pública difusa. Como si fuese posible hacer frente a la crisis de las formas clásicas de representación mediante una democracia local conectada a escala global. Hemos constatado las principales lagunas de este asunto, que se pueden reconducir a un enfoque ideológico. La idea es que las tecnologías son «neutras» por definición, y que esta presunta neutralidad es sinónimo de bondad, en tanto que fruto de una práctica de búsqueda científica objetiva, capaz de ofrecer a cada individuo lo que desea, sin esfuerzo y rápidamente.

La compleja mediación informacional puesta en marcha por Google se presenta como una membrana transparente de alta tecnología, garante de las elecciones libres de los usuarios/ciudadanos/consumidores, que usan/votan/compran sumergidos en la red libre que Google gestiona por su propio bien.

A pesar de este sueño participativo, deslumbrado por el cuento de la ciberdemocracia, pero falto de sustancia concreta, no se pueden realizar formas realmente autónomas de democracia directa mediante la centralización de las informaciones, de los saberes, de los poderes, en manos de una empresa privada (por ejemplo Google), y menos aún en las de un sujeto público (como la denominada *Authority*[1]).

1 La Internet Assigned Numbers Authority (IANA) es la Agencia de Asignación de Números de Internet. Era el antiguo registro central de los protocolos Internet, como puertos, números de protocolo y empresa, opciones y códigos. Fue sustituido en 1998 por ICANN. [N. del E.].

Ni siquiera los más avanzados de los movimientos altermundialistas escapan a la trampa de la identidad. Desean de hecho una recomposición de la identidad de clase a través de una centralidad renovada del trabajo, en este caso telemático. Pero se quedan lejos de la esfera del deseo individual incluso cuando promueven el *social networking* como solución taumatúrgica para todas las frustraciones personales, en un ritual de autoayuda tecnológica global.

Solo la autoformación puede constituir una vía de fuga a recorrer para substraerse al dominio tecnocrático. Hay que dar muchos pasos antes de poder «poner en común» algo de sí para generar sinergias. Sin una preparación técnica adecuada, el sedicente maná comunitario se revela pronto un ejercicio solipsista.

Los individuos que crean las redes tienen que aprender a gestionar sus propios datos sensibles y empezar a establecer qué hacer público y qué mantener privado. Además, deben ser capaces de decidir cuáles informaciones son «verdaderas» y cuáles son «falsas», evaluándolas en base al propio punto de vista subjetivo. En efecto, mientras recorren la Web, tienen que ser conscientes de que modifican el paisaje informacional y de que cambian ellos mismos en el curso de la navegación. Así los individuos pueden desarrollar su propia autonomía, o sea, elaborar unas reglas para cruzar los territorios virtuales y adoptar una posición personal.

Igual que cualquier otra tecnología, las tecnologías digitales no son ni buenas ni malas en sí, y como hemos visto ni siquiera son neutrales: depende del uso que se hace de ellas y del método con el que se desarrollan. Desde luego, siendo híbridos capaces de amplificar algunos aspectos de la vida real, permiten desvelar las contradicciones entre la «naturaleza» y la «cultura».

Hay que conjurar otro peligro, pues: la imagen de la Web como experiencia desmaterializada, carente de concreción física, que conduce a menudo a un rechazo luddista y reactivo de la innovación.

En esta perspectiva, la realidad virtual del ciberespacio bulliría de interacciones poco significativas, operadas por personalidades online sin memoria de las disparidades materiales de la vida concreta: sexo, raza, riqueza, posición social serían apartadas en la navegación fluida y sin fricciones de identidades ficticias.

Esta idea —solo aparentemente materialista— es la que sostienen intelectuales y observadores que, desde lo alto de sus cátedras, opinan sin haber tenido nunca la humildad de preguntar a los «niños» nacidos a la vez que las tecnologías digitales.

Por el contrario, la realidad virtual es física hasta el punto de que no podría existir sin las máquinas mecánicas, de silicio y circuitos, y sin las máquinas biológicas, es decir, los usuarios: conectado, múltiple y desterritorializado no son sinónimos de «inmaterial». Además, una actitud semejante oculta un profundo miedo a ver cambiar el mundo sin conseguir encontrar el lugar propio, y una desconfianza radical en las posibilidades de transformación y crecimiento de los individuos.

Las tecnologías digitales representan, pues, una posibilidad de liberación solo si son vinculadas con el desarrollo de álter egos digitales complejos, conscientes, capaces de interactuar de forma imprevisible. Es posible utilizar múltiples lenguajes para construir un terreno de encuentro: entre otros, Ippolita ha considerado indispensable el método experimental de la investigación científica, la riqueza inagotable de la tradición humanista, la fuerza dialógica de las pasiones políticas, la moralidad de la narración del método feminista, la curiosidad sin límites típica de la actitud del hacker. La confianza en la posibilidad de modelar tecnologías en base a los deseos de los individuos es indispensable para crear redes de verdad libres, ya sean digitales o no.

El caos de los mensajes contradictorios, el ruido de fondo a veces insoportables y la amplitud casi inconcebible de la Red pueden sin lugar a duda provocar temor, pero la exploración acaba de empezar.

Postfacio
Una historia simple

En 2006, entre los deseos de Ippolita estaba el de escribir un libro sobre el salto de la epistemología a la ontología en los mundos digitales. Nos parecía evidente que el «qué» (lo que conoces) había sido sustituido por el «quién» (lo que eres): la gestión del conocimiento se convertía en gestión y construcción de la identidad. Como no podríamos escribir un libro así (¡aunque si alguien se atreviese, le rogaríamos que nos llamara, por favor, pues estaríamos encantados de echarle una mano!) y como tampoco encontraríamos un público para el mismo —ya que ni somos académicos ni tampoco tenemos «queridos compañeros» que se interesen por nuestras circunvoluciones mentales...—, Ippolita decidió ocuparse de Google.

Un actor tan enorme que no podía pasar inadvertido. Pero todo lo que vale para Google, vale también para los demás colosos entre los motores de búsqueda (Yahoo!, Bing!, etc.). Google simplemente es el más conocido, utilizado y versátil, por lo tanto es oportuno dirigirle una dura crítica. Sin olvidar, de todas formas, que para Ippolita se trataba, y con más razón se trata ahora, solo de un modelo claro y com-

prensible para todos los usuarios de la Web preocupados por las consecuencias derivadas de la actual estructura de Internet. En este caso, de la difusión viral de comportamientos y prácticas de delegación de nuestras propias «intenciones de búsqueda», de nuestros deseos respecto a estas tecnologías, a un sujeto con miras declaradamente hegemónicas.

Ha pasado algún tiempo, el libro ha sido leído, ha sido descargado, criticado, traducido. Para nuestra sorpresa, en lugar de desinflarse frente a algún nuevo actor tecnológico, Google ha estado cada vez más en la palestra. Junto con el *cloud computing*, las «nubes» de datos, se difunde también la *FOG*, no en el sentido de niebla sino más bien de miedo: *Fear of Google*[1]. Miedo a que un monopolio del saber constituya una amenaza no solo para los individuos (que siempre están en el vagón de cola de las agendas políticas), sino también para empresas y autoridades de Estados; es decir, para el business y los poderes constituidos —en otros momentos se hubiera dicho: para el complejo militar-industrial—. Es por eso que florecen estudios y debates, se mueven las grandes corporaciones con sus pandillas de abogados para impedir que Google cambie el estatus quo. O, por lo menos, para reclamar su parte de la tarta.

Ippolita sigue sin ser un experto en el tema: la experiencia y los conocimientos se han ido construyendo a lo largo de la investigación, y no nos pasamos el tiempo estudiando los movimientos de BigG. Sin embargo ha llegado el momento de hacer un repaso de las novedades y de comprobar si el modelo de análisis propuesto continúa siendo válido. Nosotros creemos que sí lo es, porque aunque aparentemente se han producido muchas novedades, mirándolo bien todo cambia

1 FOG el «miedo a Google» hace ya tiempo que existe y se extiende desde Silicon Valley: http://www.sfgate.com/cgi-bin/article.cgi?f=/c/a/2007/05/11/MNGRIPPB2N1.DTL; y el miedo no para, sino que se sigue propagando desde Italia (http://www.corriere.it/scienze_e_tecnologie/09_ottobre_29/super-google-mondo-gaggi_4daff0a4-c45a-11de-ae8c-00144f02aabc.shtml) hasta Australia (http://www.smh.com.au/business/google-monster-20090814-el49.html).

para que todo siga igual: Google es un gatopardo contemporáneo, que persigue tenaz y eficazmente su objetivo hegemónico de «organizar el conocimiento» del mundo entero.

Google es la punta de lanza, el estándar de hecho de la «webización de todo». Toda información tiende a volverse más fácilmente accesible a través de un terminal conectado a la Web, que cada vez rima más con Google. Google, frente al caos de los motores de búsqueda llenos de publicidad invasiva, banners, etc., ha sido una gran novedad. La sobriedad y la eficiencia que lo han hecho famoso, además de la filosofía de la excelencia (en primer lugar posible sobre todo por la explotación de mucho código abierto), han llevado a académicos y tecnófilos a loarle como la alternativa *cool* en la Web, la alternativa «buena» frente a las multinacionales «malas». Tal vez el problema de fondo haya existido desde el principio: una compañía que tiene como objetivo la catalogación de todas las informaciones digitales del mundo no puede no tener objetivos hegemónicos, y es esta hegemonía alcanzada en muchos sectores lo que hoy le hace aparecer como un «peligro». Saber es poder, y si el saber significa la posesión y el acceso a los datos generados por los usuarios, está claro que quien posee las nubes de datos está en una posición de dominio. Poniéndose como punto de acceso único (o casi; privilegiado, sin duda) a la Red, Google constituye una concentración de poder inédita, con la cual los individuos particulares, las sociedades, las autoridades y cualquiera que utilice instrumentos digitales conectados entre ellos no puede evitar confrontarse.

¿Ha cambiado algo?

¿Acaso algo ha cambiado? No mucho, en realidad. Aun así vamos a elaborar un listado de las novedades de BigG —solo las que nos han sorprendido, se entiende—. Una panorámica de las «directrices evolutivas» de la máquina Google debe empezar por la apertura hacia Smartphone y los móviles de

última generación: el sistema operativo Android es ahora una realidad comercial en rápido desarrollo, una competencia temible para Windows Mobile, Iphone-Ipad, Symbian-Maemo-Meego[2]. El navegador se convierte cada vez más en un ámbito de desarrollo, como hemos recordado en otras ocasiones[3]; así tenemos inevitablemente el navegador *made in* Mountain View, Chrome, y su hermanito para open source, Chromium[4]. Puesto que un porcentaje cada vez más relevante del tiempo de uso de PC y Smartphone implica una conexión a Internet, el navegador como herramienta para navegar en el mar de la Red se hace imprescindible[5]. Una ojeada a los servicios: Google Health permite ordenar, buscar, compartir los datos médicos personales. El desarrollo y la integración con *mashups* cada vez más avanzados prosiguen rápidamente; espectaculares sobre todo los progresos en el mapeo del territorio, con Street View, que a estas alturas poco tiene que envidiarle a los sistemas de espionaje militar. Google se ha convertido también en *DNS provider* (proveedor de dominios) con Google Apps: es posible cambiar de dominio, correos incluidos, y usar el Google Calendar, GDocs y Google Sites.

En la actualidad, se ofrecen hasta 25 mailboxs de Gmail por dominio. Una bendición para las pequeñas empresas,

2 Android es una máquina virtual montada en un kernel linux modificado, o sea, construida a partir de código abierto (http://www.android.com/:). Pero, aunque Android es linux como el OS Maemo (Nokia), es más difícil de desarrollar aplicaciones en él, además de que los softwares permanecen propiedad de Google. Para conocer más de Maemo, que desde el acuerdo Nokia-Intel se denomina Meego, podéis consultar las siguientes páginas: http://maemo.org/, http://meego.com/.

3 Véase el apartado «Los navegadores como ambientes de desarrollo» (cap. 5).

4 Chrome: http://www.google.com/chrome; Chromium: http://code.google.com/chromium/.

5 No es casual que Google financie la Mozilla Foundation con 60 millones de dólares al año. ¿Qué pasaría si se le quitase la financiación, una vez puesto a punto «el navegador más rápido de todos los tiempos»? Ver http://www.zdnet.com/blog/btl/mozilla-still-too-dependent-on-google-for-revenue-can-it-diversify/27670.

nadie puede ofrecer tanto a cambio de un poco de publicidad contextual. En breve, el enfrentamiento con Microsoft y Apple llegará a su culminación con el lanzamiento de Google OS, un sistema operativo completo, basado previsiblemente en un código GNU/Linux; el navegador Chrome será sin lugar a duda el corazón pulsante de este sistema-terminal para conectarse a la «nube de Google». Además, Google no quiere quedarse fuera de las redes sociales: después de haberlo intentado con Orkut (muy difundida en Brasil), es el turno de Google Wawe y sobre todo de Buzz, que hace realidad la convergencia entre chats, galerías de foto, discusiones, RSS personales, mails y microblogging. Sobre todas estas novedades pende la hipoteca de la privacidad. No se oye hablar de otra cosa. Google representa una amenaza para nuestra privacidad, se nos repite. ¿Pero es que los demás proveedores de acceso y servicios no son una amenaza? ¿Acaso no era Google el gigante bueno?

El falso problema de la privacidad y la realidad del *profiling*

«Privacidad común» es un oxímoron, pero da buena cuenta del desconcierto que se siente descubriendo (hecho conocido desde hace años y por supuesto no descubierto por Ippolita) que existen programas de filtrado de los correos capaces de extrapolar palabras clave para proponer publicidad contextual y selectiva, sobre la base de intercambios considerados «privados, reservados». Google ha dejado patente la existencia de esferas ni públicas ni privadas, gobernadas por las tecnocracias. ¿La privacidad protege los datos sanitarios en poder de las estructuras sanitarias? ¿La privacidad protege los datos penales en mano de las estructuras judiciales? Cuando estos archivos de datos se hacen accesibles (más accesibles de cuanto lo han sido hasta ahora, porque los archivos digitales a los que se accede mediante los motores de

búsqueda son generalmente más accesibles que los archivos reales[6]) surgen los problemas. La privacidad (o, mejor aún, confidencialidad) es el «derecho a estar solos» (*the right to be alone*), e incluye todo lo que se considera privado (intimidad, pudor) y todo lo que arbitrariamente queramos mantener en privado. Pero la privacidad es una quimera. Es un concepto vacuo y poco útil, sobre todo porque comienza a existir solo cuando nos damos cuenta de que ha sido violada. Además, si por privacidad se entiende la «protección de los datos sensibles de los individuos de la mirada indiscreta de personas no autorizadas», después de Echelon (un sistema de espionaje militar e industrial de todas la comunicaciones digitales del mundo, activo desde decenios y aún hoy en perfecta forma) es evidente que estamos hablando de algo que no existe, o que no existe ya desde hace mucho tiempo.

Por otra parte, el problema no es tanto el hecho de estar bajo control: todas las llamadas son grabadas (en teoría por un tiempo limitado), todos los movimientos online registrados por alguna máquina (la de quien ofrece el acceso, por ejemplo). Sin embargo, el problema es complejo. El concepto de perfilación tal vez sea más útil: pues a Google no le interesa el individuo en cuanto a tal, sino su perfil, para comprender qué «tipo» de «cliente» potencial es para la publicidad de terceros o qué tipo de usuario/desarrollador/prosumer puede ser para sus servicios. Desde luego, es inquietante que «perfilación» sea un término procedente de las actividades del *criminal profiling* (perfil criminológico): todos fichados, todos perfilados. Pero aún más inquietante es el hecho de que a las personas, en general, no les preocupa ser objeto de perfilación, incluso quieren ser controladas de forma continuada y cada vez más invasiva, ya que son muchos (no solo los ciberfetichistas declarados) los que

6 ¿Quiénes de vosotros ha intentado consultar un fichero judicial? Pocos o nadie. Sin embargo, las informaciones sobre pedófilos y similares, al estilo «pon el monstruo en primera página», son fácilmente asequibles, especialmente en los EE.UU.

quieren hacerse con instrumentos que no saben gestionar de forma autónoma y que son asimismo instrumentos de autodelación desde abajo. No es necesario intentar fichar a las personas usando recursos sofisticados; se puede conseguir lo mismo dejando el esfuerzo directamente a los interesados: es suficiente con poner a la venta terminales evolucionados, constantemente conectados online a servicios de *social networking* (registro de las relaciones interpersonales), con videocámaras y GPS... ¡como los Smartphone y otros modelos parecidos que arrasan ahora!

En cuanto al hecho de que las fuerzas de policía, militares, servicios secretos, autoridades de distinto tipo, organizaciones criminales de todo género puedan eventualmente acceder a esos datos, el caso reciente de China resulta suficientemente ilustrativo: las autoridades chinas han forzado o hecho que se violaran las direcciones de correo electrónico de Gmail pertenecientes a disidentes hostiles al régimen para poder reprimirles. Ninguna sorpresa: una gran concentración de poder apunta casi siempre a favorecer una ansiedad irrefrenable de dominio[7].

¿Qué ha pasado?, ¿por qué el nivel de seguridad percibido está cayendo en picado?

Paranoia securitaria y militarización tecnocrática

El debate sobre la seguridad —ya sea real, presunta o percibida— se alimenta de cifras. Pero es inútil producir estadísticas: se trata de secuencias de números útiles solamente para sostener la propia hipótesis, para justificar sin explicar. Sería más bien interesante intentar comprender la génesis de

7 Más información sobre perfilación en los artículos dedicados al tema en http://www.delirandom.net, en especial: http://www.delirandom.net/#/20080321/vecchi-attacchi-allanonimato-e-svisceramento-di-questa-occulta-profilazione/ y http://www.delirandom.net/#/20091007/lip-informazionecontrollosicurezzainternet/

esta paranoia securitaria y su relación con las transformaciones provocadas por la introducción masiva de las tecnologías digitales. No es este el lugar para tratar en profundidad un tema tan complejo. Sin embargo, constatamos con placer la aparición de distintas investigaciones sobre la construcción de las políticas del miedo por parte de las autoridades[8].

En el contexto de esta investigación, esbozando una esque-matización extrema, podemos decir que las superposiciones y contaminaciones entre público/privado, opiniones subjetivas/verdades objetivas, individual/colectivo, etc., son muy difíciles de gestionar.

En lugar de aumentar las ocasiones de concertación pública para la gestión del poder tecnológico, la delegación progresiva en el uso inconsciente de los instrumentos digitales aumenta el desequilibrio tecnocrático. Google es la punta de lanza de este dominio tecnocrático en vías de expansión, hasta el punto de englobar cada aspecto de la vida cotidiana: no solo buscar informaciones, sino hablar, desplazarse, divertirse, conversar... La percepción de la fragilidad, de la exposición extrema a la que estamos sometidos por una tecnocracia despersonalizada incrementa la demanda de seguridad. De ahí el aumento de control tecnocrático, hasta llegar a una proliferación inédita de la militarización securitaria. Todo el mundo tiene que contribuir: denunciad a vuestro vecino, tiene un aire sospechoso, podría ser un terrorista, un peligro. Hay que aumentar los controles... Qué pena que el control sea justamente lo contrario de la posibilidad de confidencialidad, sin considerar el hecho de que nadie tiene la más remota idea de quién controlará a los controladores. Lo único cierto, en esta carrera por el control, es que ya ahora la cantidad de datos almacenada por los sistemas de control activos ha superado con creces las posibilidades de gestión por una parte de las autoridades (privadas y/o públicas, civiles y/o militares) que han instalado

8 Por ejemplo, CAE, *Marching Plague*, Autonomedia, Nueva York, 2006; Loretta Napoleoni y Bee J. Ronald, *I numeri del terrore. Perché non dobbiamo avere paura*, Saggiatore, Milán, 2008.

dichos sistemas. Por lo tanto, es necesario reclutar nuevos controladores que discriminen los millones de sitos indexados, las miles de horas grabadas por las videocámaras de circuito cerrado, etc., a la búsqueda de la anomalía, del peligro para la «seguridad pública». Podemos imaginar que pronto los ciudadanos serán invitados, tal vez bajo la promesa de premios en metálico, a convertirse en policías y detectives, delatores y espías. Por otro lado, el uso masivo de *spin-doctors* (responsables de imagen) en política, junto con técnicas de guerrilla de la información (*infowar*) para orientar y manipular a la opinión pública no hace sino agravar la confusión general. Es difícil orientarse en la sociedad del espectáculo en tiempo real[9].

Perspectivas e involuciones

Después de la construcción de una nueva infraestructura, después de la puesta en marcha de una red inmensa, la red tiene que funcionar. La acumulación tiene que dar sus frutos. Podemos esperar pues el lanzamiento de nuevas iniciativas, gratuitas o de todos modos de muy bajo coste (por lo menos al principio), para explotar mejor las potencialidades infinitas de los *datacenters*, de las infraestructuras de la Red y de los datos recogidos. En este sentido, el «miedo» a que Google cambie modelos enteros de negocio no es gratuito, desde la edición al periodismo y hasta la telefonía. Un ejemplo plausible: Google Provider, Google Telco, Google servicios de localización GPS 2D y 3D, Google Ediciones, Google TV/cine HD, Google espacio para MPG juegos multijugador online, Google

9 El asunto es espinoso y difícil, y a menudo los autores que se ocupan del mismo parecen padecer manías conspiracionistas, además de una curiosa oscilación entre la reivindicación de la libertad absoluta de expresión y la invocación a una autoridad garante de la verdad. Con la debida cautela escéptica, la obra del teórico de los sistemas Bruno Lussato es sin lugar a dudas un espejo inquietante de la desinformación difundida. Véase en particular Bruno Lussato, *Virus. Huit leçons sur la désinformation*, Syrtes, París, 2007, y su sito http://www.brunolussato.com.

Telemedicina, etc. En general, previsiblemente serán explorados todos aquellos servicios que necesitan de uno o más de los fundamentos a disposición de BigG; es decir, cantidad de datos en constante aumento, búsquedas personalizadas con publicidad contextual, banda ancha y espacio de disco.

Sin pretensión alguna de abarcarlo todo, mencionaremos algunas perspectivas abiertas. Google es un ejemplo meridiano de centralización y verticalidad en la organización de los mundos digitales. Sería interesante explorar y construir, no obstante, sistemas descentralizados y horizontales a partir de motores de búsqueda distribuidos, P2P. Un prototipo que ya funciona es YACY[10], que propone una especie de «pluralismo de las fuentes» en el que el *crawler* es lanzado voluntariamente por cada usuario sobre la porción de datos que desea indexar, y los resultados quedan así disponibles de forma distribuida.

El caso de China despierta numerosos interrogantes. Desde luego, los problemas que Google encuentra y encontrará en su penetración capilar a nivel global proceden del hecho de que se halla en una situación inédita, detentando una cantidad y cualidad infinita de datos sensibles y ultrasensibles (comunicaciones personales, etc.). Actualmente Google posee más datos que un Estado sobre los ciudadanos del mismo Estado... Es normal que quien ejerce el monopolio de la violencia (el Estado y sus agencias) quiera arrebatárselos. Tanto más si se trata de un Estado autoritario como es el caso de China, evidentemente sin ninguna garantía de que los utilizará mejor o peor que el mismo Google. No habiendo antecedentes, cada país reacciona a su manera... Pero la cuestión también tiene un lado técnico: ¿se puede crackear Google? Por lo visto, sí, ya que, según declaraciones de los responsables de Mountain View, las violaciones de los mailboxs se han llevado a cabo usando debilidades del sistema, y no simplemente robando o extorsionando passwords o similares.

10 http://yacy.net/.

Estos debates sobre el encuentro-desencuentro entre una empresa privada y Estados-naciones enteros abren otros escenarios inquietantes, sobre todo sobre la posible colaboración entre Google y los servicios de inteligencia. También porque, si Google fuese considerado de veras como un servicio imprescindible, estaría en la misma situación que los bancos y las demás empresas vitales para el mantenimiento del estatus quo, o sea, *too big to fail*. Si durante el desplome financiero provocado por la crisis —que desde luego, no será la última del capitalismo— ya fue muy difícil justificar, en plena era neoliberal, los rescates de bancos por parte del Estado, ¿cómo va a ser posible hacer «pública» (en el sentido de estatal) una empresa multinacional? Sin embargo, algunas reflexiones apuntan en esa dirección[11], la cual no parece muy lejos del fantasma del Nuevo Orden Mundial. Una vez más, esta inédita concentración de poder se concreta en una evidente posibilidad de dominio.

Alternativas

A menudo se cita a Ippolita como experta en Google; inevitablemente, la constatación final es: muy bien, está claro que existen una serie de problemas. Pero nosotros no somos tan expertos como vosotros. Somos simples usuarios. Somos víctimas pasivas de Google, que creíamos que era bueno y que ha resultado ser malo. ¿Cuál es la alternativa?

Las alternativas hay que construirlas. No existe una alternativa única, y en todo caso no existe ninguna alternativa si queremos un motor de búsqueda «instantáneo», un webmail de muchos gigas y además «gratuito», servicios cada vez más novedosos, velocidad cada vez mayor, todo esto sin preocuparnos de quién, cómo y para qué gestiona esos servicios. Si deseamos dar respuesta a una necesidad que crece sin límí-

11 Véase Yann Moulier-Boutang, *Multitudes*, 2009.

tes, la mejor alternativa a Google es otro Google capaz de un crecimiento aún más rápido e ilimitado.

Pero las personas podrían construir alternativas válidas de forma autónoma. La condición necesaria, y no suficiente, es la de elaborar los deseos personales antes de utilizar las tecnologías; sobre todo si el uso corriente se realiza sin ninguna conciencia, solo porque se trata de tecnologías disponibles, aparentemente gratuitas e indispensables para estar al día de las modas pasajeras; y alineados a la *dromocracia* en expansión, al culto a la velocidad del que habla desde hace tiempo Paul Virilio[12]. Por otra parte, la alternativa a un sistema tecnológico industrial y consumista, del que Google es simplemente el ejemplo más evidente, es la puesta a punto de instrumentos tecnológicos convivenciales, o sea, construidos a la medida de las personas que los utilizan, construidos por esas mismas personas junto con otras afines. Es decir, que se podría pensar en cómo se hacen las cosas (también en la Web) antes de pensar en «qué» cosas hacer[13]. Muchos ya lo hacen, difundiendo y multiplicando soluciones personales, temporales, a medida, autónomas y autodeterminadas, traducibles y modificables, placenteras, para satisfacer las propias necesidades cognitivas, los propios deseos a través de los instrumentos tecnológicos. La idea de la convivencialidad no es nueva (véase la entrada «Ivan Illich»), ha sido y continúa siendo practicada de distintas maneras, también en la Web (hay muchas herramientas de este tipo, solo hay que buscarlas... con o sin Google), y solo espera a ser desarrollada de nuevas maneras por quienes usan todos los días los instrumentos digitales.

12 Paul Virilio, *Vitesse et politique*, París, Editions galilée, 1977 (en castellano en: La Marca, Buenos Aires, 2009). Une pequeña introducción online en: http://post-media.info/post-ressources/virilio-dromologie.pdf

13 Intervención de Ippolita en la Society of the Query Conference, Amsterdam, 2009 (http://interno.eleuthera.it/webs/karlessi/finalstringato_ippolita_amsterdamSOQ_2009k.pdf), sobre tecnología convivencial; Carlo Milani, *Scritture conviviali: tecnologie per partecipare*, Milán, 2009, http://interno.eleuthera.it/webs/karlessi/tuttotesinotepiedek.pdf.

Apéndices

I. El fin del mundo en una galleta

La fecha de caducidad de Google es el 17 de enero de 2038. Esta fecha es interesante porque es precisamente dos días antes del fin del mundo. El 19 de enero de 2038 es una fecha esencial para todo el mundo POSIX, que comprende todos los sistemas operativos procedentes del sistema UNIX (GNU/Linux, BSD, Solares, etc.) y todos los protocolos de red UNIX Style (http, ftp, etc.). En los sistemas POSIX el tiempo está regulado por razones históricas por el número de segundos transcurridos desde el primero de enero de 1970. La unidad de medida para este número es un «entero con signo», valor que en informática ocupa 32 bits de memoria.

Si un programador crea una variable de tipología «entero con signo» para memorizar un valor numérico, este puede ser como mínimo –2147483648 y como máximo 2147463647. Un número muy grande, pero que se convierte en un valor muy pequeño si lo transformamos en segundos. En 32 bits solo caben apenas 136 años.

Cuando Ken Thompson, casi por juego, inventó Unix, nunca hubiera imaginado que iba a revolucionar tanto la informática y, sobre todo, nunca hubiera pensado que una serie de sus elecciones (el filesystem, la hora, la codificación de los caracteres, las llamadas a función, el lenguaje de programación, etc.) iban a convertirse en ley para todo programador del globo en muy poco tiempo. Thompson decidió en 1972 que la fecha 0 de su sistema operativo sería el 1 de enero de 1970.

2147463647 segundos después del 1 de enero de 1970, será exactamente el 19 de enero de 2038 a las 3:14 de la mañana, último segundo del mundo Unix, de Internet (que funciona

gracias a protocolos Unix). Después de ese segundo nos veremos proyectados al 13 de diciembre de 1901 a las 20:45. Seguramente éste será un problema que tendrá que ser gestionado en los próximos años y requerirá un cambio que marcará una nueva época en la gestión del tiempo y de todo lo demás en los sistemas Unix. Sin duda un problema algo más serio y fundamentado que el tan cacareado *millennium bug*, el «efecto 2000», en el paso del año 1999 al 2000.

La fecha de caducidad de Google es pues la máxima fecha que se puede configurar en un *cookie* y que conlleva que el navegador de nuestros sistemas operativos no eliminará nunca aquel cookie y las informaciones allí guardadas. Para mayor información sobre el problema de las fechas véase:

– http://www.2038bug.com/index.html
– http://www.iee.org/Policy/Areas/SCS/problemdates.cfm
– http://en.wikipedia.org/wiki/Unix_epoch~~

II. Interzonas: influencias y dominios en los mundos digitales

Los proyectos open source de Google están hospedados también en el portal SourceForge.net; no hay cambios sustanciales, ni siquiera cambios gráficos relevantes, en el paso entre el sitio de google.com y el de SourceForge.net. Los distintos proyectos permanecen inalterados casi hasta en los mínimos detalles, tanto en el código como en la interfaz. El único elemento digno de destacar es el cambio del servidor anfitrión: de hecho se pasa de una sección con el nombre del proyecto en el sitio de code.google.com a una página hospedada en SourceForge.net en un dominio de tercer nivel (por ejemplo, nombreproyecto.SourceForge.net). Existen, pues, dos accesos separados e independientes para alcanzar la misma interzona. Sin embargo, estas zonas pertenecen a dominios distintos: google.com y SourceForge.net.

El concepto de dominio está en la base de la estructura de las redes. En el ámbito informático, con el término «dominio» se entiende el nombre de un server dentro de una escala jerárquica extremamente compleja gestionada por una red de ordenadores llamados DNS (Domain Name Server). Estos ordenadores se encargan de gestionar la asociación server-dominio y de permitir a todo el mundo alcanzar unas direcciones esparcidas en la Red gracias a su nombre.

Por ejemplo, cuando se quiere llegar a www.google.com, un programa específico dentro de nuestro ordenador pregunta a uno de los DNS qué server responde al nombre www.google.com. En sustancia, la red de los DNS se encarga de buscar dentro de gigantescos mapas y bases de datos el server que le pedimos, asociando la dirección alfabética que nos resulta comprensible a una dirección numérica (el número IP) que es comprensible para las máquinas. Nos comunica luego el resultado, y nos conecta a Google o a cualquier otra dirección solicitada.

Simplificando mucho este complejo mecanismo jerárquico, podemos decir que los dominios están estructurados por niveles. En una dirección web típica, los niveles se leen de derecha a izquierda. Los primeros niveles son pues los genéricos .com, .org, .net, y los nacionales como .it, .es, .fr, .de, .tv. En el segundo nivel hallamos los nombres reales de los server, asociados a uno solo de los dominios de primer nivel: google.com, eleuthera.it, somasuite.net, alekoslab.org., etc. Nada le impide a un server tener más dominios de segundo nivel ligados a distintos dominios de primer nivel: google.it, google.de, google.com son todos dominios de segundo nivel válidos. No es obligatorio que estos dominios de segundo nivel se refieran a una misma entidad: por ejemplo, ippolita.net es el sitio del proyecto ippolita, ippolita.it es el sitio de la poetisa Ippolita Avalli, e Ippolita.com es el sitio de Ippolita que fabrica y vende joyas.

Desde el segundo nivel en adelante la gestión corresponde a quien gestiona el segundo nivel. Por ejemplo, code.google.com está gestionado directamente por Google, así como www.google.com y también moon.google.com.

No existe una regla precisa que impone la forma en la que se dan los nombres de tercer nivel: pensemos por ejemplo en una dirección evocadora, pero absolutamente correcta, como del.icio.us: el sufijo .us indica que pertenece al dominio nacional USA, el nombre «icio» indica el server, el prefijo «del» de tercer nivel completa el juego de palabras, de forma que la dirección total signifique «delicioso» en inglés. Sin embargo, existe la costumbre de asociar algunos dominios a servicios específicos. Una dirección del tipo www.server.net hospedará tendencialmente la primera página del servicio de server.net, mientras mail.server.net hospedará servicios de correo electrónico.

La imagen tal vez más adecuada para describir esta jerarquía es la de una casa: los servicios tienen una puerta de acceso directo (mail.server.net) mientras que el contenido de los distintos servicios puede ser hospedado dentro de una

de las habitaciones interiores (www.server.net/fotografie/), o sea, en uno de los directorios del servidor principal.

La gestión de los dominios de tercer nivel de Google es ilógica, y por ello extremadamente abierta y fluida. De hecho es curioso notar que a la Luna se le ha concedido un dominio de tercer nivel (moon.google.com) mientras que a Marte no, relegado a un directorio secundario (www.google.com/mars): ambos servicios permiten explorar mapas, sin diferencias significativas en el servicio que proponen.

La estructura de SourceForge.net es más rígida, así como la de Freshmeat.net (que Google de momento no utiliza) y de otros proyectos análogos. SourceForge.net ofrece los proyectos hospedados como servicios para los usuarios a causa de su misma naturaleza de portal; por esta razón, todos los proyectos de Google se encuentran en un dominio de tercer nivel. Se trata de un instrumento para las comunidades de desarrollo que dedica a cada una de ellas un ambiente hecho a medida y rico en servicios. Las distintas comunidades piden los servicios que necesitan y SourceForge.net les da un espacio completo y personalizado. Por lo tanto, Google se comporta como la comunidad de Gaim o de cualquier otro software: y, sin embargo, Google no es precisamente una comunidad cualquiera de desarrolladores independientes...

Otras interzonas se crean para gestionar los campos de influencia de las correspondientes comunidades y la eventual interacción entre proyectos que de ellas dependen. El proyecto Open Source Mozilla y la página predefinida para su navegador Firefox es un pequeño ejemplo de gran visibilidad pública. Cada vez que un usuario instala el navegador Mozilla-Firefox, la primera página online que verá al empezar Internet será la página de Google modificada gráficamente de forma que incluya también el logo de Firefox.

Google facilita también *toolbars* (barras de herramientas) para que se integren en Firefox. Las *toolbars* son bloques gráficos (en general en forma de línea o columna) que, gracias a una serie de botones y otros objetos gráficos, activan

funciones particulares del programa que las hospeda. Las extensiones para Firefox permiten añadir barras de herramientas gratuitas dentro del propio navegador. Lo que el público difícilmente sabe es que tales productos no son en absoluto open source. En efecto, la *toolbar* de Google no es un producto descargable de los clásicos sitios de extensiones para Mozilla, porque está sujeta a una licencia propietaria. El funcionamiento de la barra de herramientas está oculto dentro de una pequeña librería usada solo por este instrumento y disponible sin código fuente. Existen distintos estudios ilegales de *reverse engeneering* (ingeniería inversa) que muestran el funcionamiento de esta extensión y la forma en la que se solicitan datos del sitio de Google.

La integración entre el propio terminal (ordenador, laptop, etc.) y el Google-server es cada vez más estrecha; de hecho, una de las novedades entre las extensiones para Mozilla Firefox se llama Google Browser Sync: *bookmarks*, cronologías, *cookies*, passwords guardados e incluso fichas y ventanas dejadas abiertas durante la última sesión, se guardan online en los servidores de Google y luego son sincronizadas automáticamente en las distintas ubicaciones usadas y en las nuevas sesiones de navegación. El control sobre el usuario se extiende a cualquier dispositivo capaz de acceder a la Web.

Por último, podemos identificar otra interzona en el frente de las licencias de uso. Google ha intentado adecuarse al estándar de las demás empresas open source que tienden a utilizar versiones más o menos retocadas de la BSD 2.0 (a excepción de la Novell). Esta licencia permite compartir el código pero también un fuerte control sobre los programas y sus versiones. Por un lado, esto permite la creación de comunidades, mientras que, por otro, introduce con todos los medios al alcance control general a lo largo de todo el proceso. Cualquier programador independiente que desarrolle gratuitamente, utilizando esta licencia, podrá ser legalmente expropiado en cualquier momento de su propia creación.

Por mucho que algunos responsable de Google hayan comunicado un posible cambio de licencia a favor de la licencia Apache, MIT o incluso GPL, la BSD sigue siendo la licencia estándar para todos los proyectos hospedados y promovidos por el motor de búsqueda. BSD constituye, por lo tanto, la interzona de patente útil para atraer hacia sí los códigos libres del mundo open source y hacer de ellos propiedad exclusiva de la propia empresa.

Ramón Fernández Durán

Tercera Piel

Sociedad de la imagen y conquista del alma

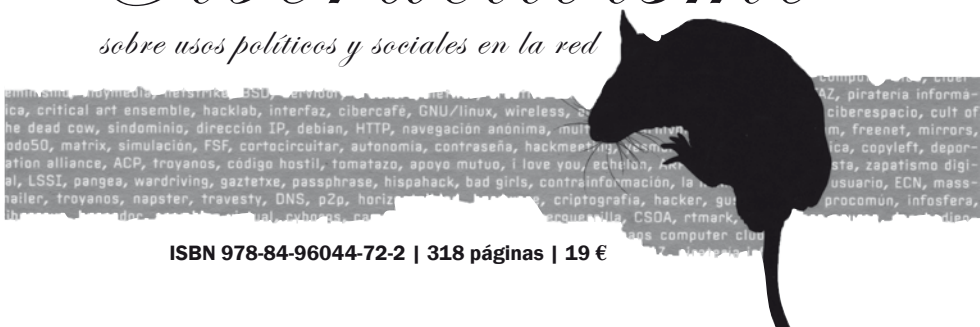
Allí donde en muchas ocasiones no llega el agua, llega la televisión. La irrupción de los medios de comunicación de masas liderados por ésta, y de las nuevas tecnologías de la información y la comunicación (NITC), ha impulsado el desarrollo de la sociedad de consumo (mediante la constante generación de nuevas y falsas necesidades) y, además, ha permitido crear las condiciones para proyectar fuera de los espacios centrales los valores urbano-metropolitanos, propiciando una capacidad de imposición mundial sin precedentes de los valores e intereses dominantes de Occidente.

ISBN 978-84-92559-19-0 | 80 páginas | 5 €

Reunión de Ovejas Eléctricas (roe)

Ciberactivismo

sobre usos políticos y sociales en la red


ica, critical art ensemble, hacklab, interfaz, cibercafé, GNU/linux, wireless,
ciberespacio, cult of
he dead cow, sindominio, dirección IP, debian, HTTP, navegación anónima,
m, freenet, mirrors,
odo50, matrix, simulación, FSF, cortocircuitar, autonomía, contraseña,
ica, copyleft, depor-
ation alliance, ACP, troyanos, código hostil, tomatazo, apoyo mutuo, i love you, echelon,
sta, zapatismo digi-
al, LSSI, pangea, wardriving, gaztetxe, passphrase, hispahack, bad girls, contrainformación,
usuario, ECN, mass-
ailer, troyanos, napster, travesty, DNS, p2p,
procomún, infosfera,
CSOA, rtmark,

ISBN 978-84-96044-72-2 | 318 páginas | 19 €